

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

QM protein - protein search, using sw model

Run on: November 12, 2004, 12:54:14 ; Search time 42 Seconds
(without alignments)
1466.160 Million cell updates/sec

Title: US-10-730-010-2

Perfect score: 3404

Sequence: 1 MDKYDVKAIGQAGFKAYL.....LCANDCSLKDSEKEMELRT 640

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : PIR 79:*

1: pir1:*

2: pir2:*

3: pir3:*

4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	1078	31.7	774	2 S25284	protein kinase nek
2	659.5	19.4	792	2 JC7122	protein kinase (EC
3	655	19.2	841	1 I78885	serine/threonine-s
4	569	16.7	357	2 T29771	hypothetical prote
5	535.5	15.7	445	2 G01452	NIMA-like protein
6	520.5	15.3	431	2 T11854	protein kinase (EC
7	512.5	15.1	699	2 A43734	probable protein k
8	507.5	14.9	722	2 T37970	probable G2-specif
9	471.5	13.9	338	2 JC7838	Nek6 protein kinas
10	459	13.5	779	2 A57177	NIMA-like protein
11	451	13.2	941	2 T49136	protein kinase-lik
12	424	12.5	294	2 T21075	hypothetical prote
13	423	12.4	435	2 T23580	probable protein k
14	415.5	12.2	200	2 B96587	hypothetical prote
15	410	12.0	648	2 T47988	serine/threonine-p
16	405.5	11.9	1895	2 T15881	hypothetical prote
17	396.5	11.6	114	2 T38224	protein-serine/thr
18	395.5	11.6	1142	2 S53359	GIN4 protein - yea
19	394	11.6	925	2 A5748	protein kinase (EC
20	393.5	11.6	576	2 S22127	protein kinase pol
21	393.5	11.6	705	2 A48144	protein kinase CDC
22	393	11.5	465	2 B55748	protein kinase (EC
23	391	11.5	658	2 T39500	serine/threonine-s
24	387.5	11.4	1051	1 JMW051	serine/threonine-s
25	385	11.3	1558	2 T23253	hypothetical prote
26	384.5	11.3	733	2 A57459	ribosomal protein k
27	381	11.2	257	2 G84797	probable protein k
28	381	11.2	461	2 T14822	probable serine/thr
29	379	11.1	445	2 T50802	serine/threonine p

30 379 11.1 682 2 A44493 serum-inducible ki
31 377 11.1 490 2 S47946 protein kinase hom
32 376 11.0 915 2 S74283 probable protein k
33 375 11.0 819 2 A53714 protein kinase (EC
34 374.5 11.0 752 1 A32571 ribosomal protein
35 373.5 11.0 608 2 G96575 probable MEK kinas
36 373.5 11.0 856 2 T43631 serine/threonine k
37 372.5 10.9 735 2 T51901 ribosomal protein
38 372.5 10.9 740 2 T24340 hypothetical prote
39 372.5 10.9 797 2 T23927 hypothetical prote
40 372.5 10.9 897 2 S61137 probable membrane
41 371.5 10.9 363 2 T08542 mitogen-activated
42 371.5 10.9 363 2 T51735 mitogen-activated
43 370 10.9 982 2 T18576 serine-threonine k
44 369 10.8 1062 2 S46367 protein kinase CDC
45 368.5 10.8 795 2 JC4234 gene fused protein

ALIGNMENTS

RESULT 1

S25284

protein kinase nek1 (EC 2.7.1.1) - mouse

C:Species: Mus musculus (house mouse)

C>Date: 17-Apr-1993 #sequence_revision 17-Apr-1993 #text_change 16-Aug-2004

C:Accession: S25284

R:Letwin, K.; Mizzen, L.; Motro, B.; Ben-David, Y.; Bernstein, A.; Pawson, T.

EMBO J. 11, 3521-3531, 1992

A:Title: A mammalian dual specificity protein kinase, Nek1, is related to the NIMA cell

A:Reference number: S25284; MUID:93010942; PMID:1382974

A:Accession: S25284

A:Status: not compared with conceptual translation

A:Molecule type: mRNA

A:Residues: 1-774 <LET>

A:Cross-references: UNIPROT:P51954; GB:S45828; NID:g256854; PIDN:AAB23529.1; PID:g256855

C:Genetics:

A:Gene: nek1

C:Superfamily: protein kinase homology

C:Keywords: ATP; phosphotransferase

F:2-259/Domain: protein kinase homology <KIN>

F:10-18/Region: protein kinase ATP-binding motif

Query Match 31.7%; Score 1078; DB 2; Length 774;
Best Local Similarity 36.4%; Pred. No. 2.5e-35;
Matches 260; Conservative 107; Mismatches 171; Indels 176; Gaps 21;

Qy 1 MDKYDVKAIGQAGFKAYLAKGSKGHCVCVKEINFEKMPIOEKEASKEVILEKMKH 60

Db 1 MEKTVRLQKIGEGSGFKAVLVKSTEDGRHYVVIKEINISMSDKEROESREAVLANMKH 60

Qy 61 PNIVAFNSFOENGRFLFVMEYCDGDLMKRINRQGVLPSEDOILGWVQVISLGLKHII 120

Db 61 PNIVQYKESFEENGSLYVMDYCEGGDLFKRIQAQKALFOEQIIDLWFVQICLAKHVE 120

Qy 121 DRKILHRDIKAQNFISKNGWAKLGDGFIARVNNNSMELARTCIGTPYYLSPEICNKP 180

Db 121 DRKILHRDIKSNIFLTGQTV-QLGDFGIARVNNNSMELARTCIGTPYYLSPEICNKP 179

Qy 181 YNNKTDIWSGCVLYELCTLKHPEGNLQOLVLTQAHFAIPISCFSELSLSOLF 240

Db 180 YNNKSDIWAIGCVLYELCTLKHAFEGNKNLVKISGFPFVPSHYSDLSLSOLF 239

Qy 241 QVSPDRPSINSILKRPFLNLIPIKVLTPVEIIEEFSHMLICRAG----- 285

Db 240 KRNPRDRPSVNSILEKGFIAKRIEFLSPQLIAEEFLKTLKSKFGPQLPKRPSAQGV 299

Qy 286 -----APASRHAGKVQCKTKQVFRGKCPKPRRIS--VPIKRNAILHRNEWRP 333

Db 300 SSFVPAQKITPKAAKGVPLTYKKGDKLLEKPPPKQAQHPVKK---MNSGEERX 356

Qy 334 PAG--AQKARSIMKIER----- 348

Db 357 KQSEAAKRRLEFIEKQKQDQIRFLKAEQKQKQKORLERINRAREQGWENVLKAGG 416
QY 349 -----PKIAVCGHYDYVAAQDMLRR-RAH-----KPSYHP--IPQE 393
Db 417 SCEVKASFFGIGGAVSPSPRCQYEHYHAFQMQRLRAEDNEARWKGGIYGRMLPER 476
QY 384 NTG---VEDYQETRHGSPSQWPAEYLQKFEAQVKKLVE-----KOLGLRPSS 431
Db 477 QKGLAVERANQ-----VEEFLQKREAVQKARABGVVYLARLQIRLQ--- 522
QY 432 AEPYNQORQLR-----SNGEPRFQELPPFRKMKKEQYWKQLEBIRQOYLN 479
Db 523 ---NFNERQIQKAKLGENKEADTKGQEAT-BETDMR---LKKMESLKQAQTNARAVALK 575
QY 480 DMKIRKKGPEPEDEKDLQKQMLQNTKSKNP-----EOKYKAKGVKF--- 525
Db 576 EQLERKKEAYERE--KKVWEHULVARVKSDDVPLPLELLETGGSPSKQVQKPVISVTS 633
QY 526 --EINLQKICDENILOEEFAMDFNETLT-----FEDGMKFEYECVKEH 569
Db 634 LKEVLDGSLTD---TQEEWKEKSNASISKKRILRLNENLKAQDEKEKQH 683
RESULT 2
JC7122
protein kinase (EC 2.7.1.37) 2 - mouse (strain balb/c)
N:Alternate names: serine (threonine) protein kinase
C:Species: Mus musculus (house mouse)
C:Date: 23-Mar-2001 #sequence_revision 23-Mar-2001 #text_change 09-Jul-2004
C:Accession: JC7122
R:Hayashi, K.; Igarashi, H.; Ogawa, M.; Sakaguchi, N.
Biochem. Biophys. Res. Commun. 264, 449-456, 1999
A:Title: Activity and substrate specificity of the murine STK2 serine/threonine kinase
A:Reference number: JC7122; MUID:20001940; PMID:10529384
A:Accession: JC7122
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-792 <HAY>
A:Cross-references: UNIPROT:Q921J2; GB:AU223071; NID:94138209; PID:94138209
C:Keywords: human serine/threonine-specific protein kinase STK2; protein kinase homol
C:Keywords: phosphotransferase

Query Match 19.4%; Score 659.5; DB 2; Length 792;
Best Local Similarity 29.5%; Pred. No. 4.5e-19;
Matches 157; Conservative 108; Mismatches 161; Indels 107; Gaps 14;
QY 4 YDVVKAIGGAFGKAYLAKGSDSKHCVIKINFEKMPIOKEASKKEVILLKMKHPNI 63
Db 6 YCYMKVGRGSGYEVTLVKHRRDQGVYIKLNLNASSRRERRAAEQEQAQLLSQKHPNI 65
QY 64 VAFNFSFO-ENGRLPIVMEYCGDGLMKRINRQGVLFSEDIQLGWFWVQISLGLKHIDR 122
Db 66 VTYKESWEGDGLLYVMGFCGGDLVYKLEKQGLLPESQVWVFVQIAVALQYLHEK 125
QY 123 KILHRDIAQNIIFLSKNGWAKLGDGFIARVLNNSMELARTICIGTPYVLSPEICQKPNY 182
Db 126 HILHRDLTKQNVFLTRNII-KVGDGLIARVLEHNDMASTLIGTPYVMSPELFSNKPYN 184
QY 183 NKTIDWSLGCULYELCTLKHPEGNNLQOLVLIKCOAHFAPISPGFSRELHSLISQLFQV 242
Db 185 YKSDVWALGCCVYEWATLKHAFNAKMNSLVYRIIEGKLPMKPVYVSTELAEILRTLSR 244
QY 243 SPRDRPSINSILKRPFLNLIIPKYLTPTEVIOEFSHMLIC----- 282
Db 245 RPEERPSVRSILRQPIYKQISFFLEATKIKTKNNKNGDSQSKFPATVVSGEAESNHE 304
QY 269 ---PEVIOEFSHMLICRAG-----APASRHAGKVQK----- 298
Db 305 VIHQPPLSSEGQTYIMGEGKLSQEKPRASGLKSPASLKAHTKQDLSNTTELATISS 364
QY 299 ---CKIQV-----RFRGKPPRSR 315
Db 365 VNIDILPAKGRDSVSDGFVQENQPRYLDASNELGICISQVEEEMLDQNTKSSAQENL 424
QY 316 ISVPIKRNAIL--HNE-----WRPPAGAKARSIKMIERP----- 350
Db 425 I--PMWSSDIVTGEKNEPVKPLQPLIKEQKPKDQSLSPKLECSGTTLAHNSLRLGSS 482
QY 351 -----TAAVCGHYDYVAAQDMLRRRAHAKPSYHP-IPQENTGVEDYQETRHGP 398
K-CPPRSRISVPIKRNAILHRNEWPPAGAKARSIKMIE--RPKIAVCGHYDYVAAQ 365

Db 358 ESATISIRINIDIL-----PAERDSANAGVQESQPOHVDAADEV---SQC 402
QY 366 DMLRRRAH-----KPSYHP--IPQENT-GVEDYQETRHGSPSQWPAEYLQKFEAQ 416
Db 403 SISQEKRLQNTKSSDQFGNLLPFRSDGSGSELVKPLYPFN-----KQKPPDDQ 457
QY 417 YKLKVEKQGLRP-----SSAEPNYN-----ORQELRSNGEPRFQELPFR 457
Db 458 VTGIIENODSIHPSRQSPHSSNSSELSRQRQKKEQTAHSGTSQRCLEPPR 510
RESULT 3
I78885
serine/threonine-specific protein kinase (EC 2.7.1.1) STK2 - human
C:Species: Homo sapiens (man)
C:Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 09-Jul-2004
C:Accession: I78885
R:Revedakou, E.N.; He, M.; Baptist, E.W.; Craven, R.J.; Cance, W.G.; Welcsh, P.L.; Simmor
Oncogene 9, 1977-1988, 1994
A:Title: Two novel human serine/threonine kinases with homologies to the cell cycle regul
A:Reference number: I58396; MUID:94268838; PMID:8208544
A:Accession: I78885
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-841 <RES>
A:Cross-references: UNIPROT:P51957; GB:L20321; NID:9348244; PID:AAA36658.1; PID:9348245
C:Genetics:
A:Gene: GDB:STK2
A:Cross-references: GDB:374125
A:Map position: 3p21.1-3p21.1
C:Superfamily: human serine/threonine-specific protein kinase STK2; protein kinase homol
C:Keywords: phosphotransferase
F:4-261/Domain: protein kinase homology <XIN>
Query Match 19.2%; Score 655; DB 1; Length 841;
Best Local Similarity 24.0%; Pred. No. 7e-19;
Matches 195; Conservative 136; Mismatches 251; Indels 230; Gaps 22;
QY 4 YDVVKAIGGAFGKAYLAKGSDSKHCVIKINFEKMPIOKEASKKEVILLKMKHPNI 63
Db 6 YCYLWVGKSGYEVTLVKHRRDQGVYIKLNLNASSRRERRAAEQEQAQLLSQKHPNI 65
QY 64 VAFNFSFO-ENGRLPIVMEYCGDGLMKRINRQGVLFSEDIQLGWFWVQISLGLKHIDR 122
Db 66 VTYKESWEGDGLLYVMGFCGGDLVYKLEKQGLLPENQVWVFVQIAVALQYLHEK 125
QY 123 KILHRDIAQNIIFLSKNGWAKLGDGFIARVLNNSMELARTICIGTPYVLSPEICQKPNY 182
Db 126 HILHRDLTKQNVFLTRNII-KVGDGLIARVLEHNDMASTLIGTPYVMSPELFSNKPYN 184
QY 183 NKTIDWSLGCULYELCTLKHPEGNNLQOLVLIKCOAHFAPISPGFSRELHSLISQLFQV 242
Db 185 YKSDVWALGCCVYEWATLKHAFNAKMNSLVYRIIEGKLPAWPRDYSPELAEILRTLSK 244
QY 243 SPRDRPSINSILKRPFLNLIIPKYLTPTEVIOEFSHMLIC----- 268
Db 245 RPEERPSVRSILRQPIYKQISFFLEATKIKTKNNKNGDSQSKFPATVVSGEAESNHE 304
QY 269 ---PEVIOEFSHMLICRAG-----APASRHAGKVQK----- 298
Db 305 VIHQPPLSSEGQTYIMGEGKLSQEKPRASGLKSPASLKAHTKQDLSNTTELATISS 364
QY 299 ---CKIQV-----RFRGKPPRSR 315
Db 365 VNIDILPAKGRDSVSDGFVQENQPRYLDASNELGICISQVEEEMLDQNTKSSAQENL 424
QY 316 ISVPIKRNAIL--HNE-----WRPPAGAKARSIKMIERP----- 350
Db 425 I--PMWSSDIVTGEKNEPVKPLQPLIKEQKPKDQSLSPKLECSGTTLAHNSLRLGSS 482
QY 351 -----TAAVCGHYDYVAAQDMLRRRAHAKPSYHP-IPQENTGVEDYQETRHGP 398

Query Match 15.3%; Score 520.5; DB 2; Length 431;
Best Local Similarity 34.9%; Pred. No. 6.6e-14;
Matches 110; Conservative 60; Mismatches 114; Indels 31; Gaps 6;

QY 10 IGQAFKAYLAKGSKHCVKEINFEKMPQOEKASKEVILEKMKHPNIVAFNS 69
DB 26 VGLGYGEAYVAESVEDGSLCAVKWMLSDQRDKRYAQSEIKKLANCHNPHNIYIED 85
QY 70 FOENGRFLVMEYCDGDLKMRINRGV---LFSDDQILGWFOVLSGLKHIDRKIL 125
DB 86 HEENDRLVMEFADSGNLEQI-KLRGSGDARYFOEHEALFLQCLALDYHSHKML 144
QY 126 HRDIAQNIFLSKNGMAKGLDGFQIARVLANSE--LARTCIGTPYLYSPEICQNPYNN 183
DB 145 HRDIXSANVLLTSTGLV-KLGDGFGSHQYEDTVSGVVASTFCGTPPYLAPELWNNKRYN 203
QY 184 KTDIWSLGCVLVECTLKHPPEGNLQOLVLIKCOAHFAPISPGFRELHSLISQLFQVS 243
DB 204 KADVMSLGLVLYETMGWKKPSPASNLKGLMSKVLAGYAPDPFSSEFKRVVDGILVAD 263
QY 244 PRDRPSINSILKRPFLNLPKYL-----TPEVIQBEFSSMLICRAGAPASR- 290
DB 264 PNDPSVREIFQIPIYINKGLKLFVQALKKNERISDSVKELVTVQVSLISSEVSPDAHRF 323
QY 291 -----HAGKV 295
DB 324 LVSQINVDVTHRGHV 338

RESULT 7
A43734
Probable protein kinase nima (EC 2.7.1.-) - Emericella nidulans
C:Species: Emericella nidulans, Aspergillus nidulans
C:Date: 03-Mar-1993 #sequence_revision 03-Mar-1993 #text_change 09-Jul-2004
C:Accession: A43734; S28786
R:Osmani, S.A.; Pu, R.T.; Morris, N.R.
Cell 53, 237-244, 1988
A:Title: Mitotic induction and maintenance by overexpression of a G2-specific gene that
A:Reference number: A43734; MUID:88194523; PMID:3359487
A:Accession: A43734
A:Molecule type: DNA
A:Residues: 1-699 <OSW>
A:Cross-references: UNIPROT:P11837; GB:M20249; NID:g168065; PIDN:AAA33316.1; PID:g168066
C:Genetics:
A:Gene: nima
C:Superfamily: unassigned Ser/thr or Tyr-specific protein kinases; protein kinase homolo
C:Keywords: ATP; autophosphorylation; nucleus; phosphoprotein; phosphotransferase; serin
F:17-25/Region: protein kinase ATP-binding motif

Query Match 15.1%; Score 512.5; DB 2; Length 699;
Best Local Similarity 26.1%; Pred. No. 2e-13;
Matches 181; Conservative 101; Mismatches 231; Indels 181; Gaps 26;

QY 2 DKYDIKAGQAFKAYLAKGSDSKHCVKEINFEKMPQOEKASKEVILEKMKHP 61
DB 9 DKYEVLEKIGCGSGFIIRKVRKSDGFTLCRKEINYIKNSTKEREQLTAEFNILSLRHP 68-
QY 62 NIVAFN--SQENGRFLVMEYCDGDL-----MKRINRQGVLPSEDIQILGWV--Q 111
DB 69 NIVAYIHEHLKASODLYLMEYCGGDLMSVINKLRTNK-----YABEDFV-WRLSQ 122
QY 112 ISLGLKHIH-----DRKILHRDIAQNIFLSKNGMAK 144
DB 123 LVATALYRCHYGTDAEVGNSLLGPAPKPSGLKQKQAOQTILHRDLKPNIFLGSDNTV-X 181
QY 145 LGDFGIARVLANSEMLARTCIGTPYLYSPEICQNPYNNKTDIWSLGCVLVECTLKHPF 204
DB 182 LGDFGLSKLM-HSHDFASTYVGTFFYMSPEICAAEKYTLRSDIWAUGCIMFELCQRPFF 240
QY 205 EGNLQQLVLIKCOAHFAPISPGFRELHSLISQLFQVSPDRDRPSINSILKRPFLNLP 264

DB 241 NARTHIQLVKIRGKFAPLPDFYSSSELKNVIAASLAVNPDHPRDTATLINTPVI----- 295
QY 265 KYLTPEVIOEEFSSMLICRAGAPASRHAGKVQCKIQKVRF-----RGKPPPSRISVPI 320
DB 296 -----RLMRREVELNLSRA---ARKKEEATMQAKDVEQAFKLEKEKQOIRSELENSI 347
QY 321 KRNAILHRNWRPAGACAKARSINKIERPKTAAVCGHYDYVYAOLDMLRRRAHPSYHPI 380
DB 348 RR-----EWEVKARUEIDRQVQ-----NELDKLRKRFCEVDQV 382
QY 381 PQENTGVEDYQOETRHGSPSQWPAEYLQRFKFAQQYKLKVEKQLGLRPSAEPNYQ 440
DB 383 AQE-----VEKQ-----RRNANYREDA 399
QY 441 ELRNGEPRFOELPFRKNEMKEQYWKQLEIEIRQQYL-----NDMKIRKQGR----- 490
DB 400 SLRSSGHSQ-----SSNSDSDPSSSDISQLESPTNKAALPKKESPTPTRS 453
QY 491 -----EPEDIEK-----DLQKMLQNTKESKNPEQKYKAKGVKFEINLDKIS 534
DB 454 KTVVDSFNDIQWAPSPISIASLSLSRPTSATYSGKNIFAEGERKR-PKFEPTL--AYS 510
QY 535 DENILOEEAMDPNET--LTFEDGMKPEYECVKEHGDYTDKAFKELHCEPAF-----T 588
DB 511 DD-----EDDTPELSPTRPKVPKDPDFKAPSRPLLRQ-----TTALWQKLSTQPIFPANPS 563
QY 589 ELTWLSFLFLEYS--LPHFLLEKSPFSRHIED 619
DB 564 RLQMSAPDVRKESRSRPHRLSKIPSSANLAAD 597

RESULT 8
T37970
Probable G2-specific protein kinase (EC 2.7.1.-) - fission yeast (Schizosaccharomyces pom
C:Species: Schizosaccharomyces pombe
C:Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 16-Aug-2004
C:Accession: T37970
R:Badcock, K.; Churcher, C.M.; Barrell, B.G.; Rajandream, M.A.; Wood, V.
submitted to the EMBL Data Library, September 1997
A:Reference number: Z21758
A:Accession: T37970
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-722 <BAD>
A:Cross-references: UNIPROT:O13839; EMBL:Z98975; PIDN:CAB11653.1; GSPDB:GN00066; SPDB:SP7
A:Experimental source: strain 972h-; cosmid c19E9
C:Genetics:
A:Gene: SPDB:SPAC19E9.02
A:Map position: 1
A:Introns: 20/3; 28/3
C:Superfamily: protein kinase homology
C:Keywords: phosphotransferase; protein kinase

Query Match 14.9%; Score 507.5; DB 2; Length 722;
Best Local Similarity 25.4%; Pred. No. 3.3e-13;
Matches 176; Conservative 120; Mismatches 277; Indels 121; Gaps 25;

QY 1 MDKYDIKAGQAFKAYLAKGSDSKHCVKEINFEKMPQOEKASKEVILEKMKH 60
DB 1 MEKYKILECIGHSGFRIYKVQRLKDGALLAQKETHFGNITRQEKYIADENVILRLKH 60
QY 61 PNIVAFNSFQNGR---LFTVMEYCDGDLKMKINR--ORGVLFSDQILGWFOVLSL 114
DB 61 PNIVQCG--BELNRSQVINLYMEYCGHGLANLIQRYKEEKRFTEQVLEKFTQLLL 118
QY 115 GLKHIH-----DRKILHRDIAQNIFLSKNGMAKGLDGFQIARVNL 155
DB 119 ALYRCHYGENAPACDSQWPREIFFHPKQSVLHRDIKPNIFLDENNSV-KLGDGFLSKLLD 177
QY 156 NMEELARICIGTPYLYSPEICQNPYNNKTDIWSLGCVLVECTLKHPFEQGNLQQLV 215
DB 178 NTRVFTQSYGVTPPYMSPEIRSSPYSAKSDVWALGCVIFECMLTHPEGRSYLELQ 237

QY 216 ICAHAPAPISPGSRRLHSLISOLFQVSPDRDRPSINSILKRPFLNLIPIKYITPEVIOEE 275
Db 238 ICGNLSWDHYSDFVLLIRHCLVNSDLRTTYQLLRSPILSDIRSKLESERVVLQ 297
QY 276 FS-----HMLICRAGAPASRHAGKVQKIOKVRFCGKCPSPRSRISVPIKRNAILHRN 329
Db 298 SLLHKKHQMLIQLENDLQFREQLSARSELENV-----IASHLA---QREILIRE 347
QY 330 EWPPAGAKARIKMIERPKIAAVCGHYDYVYAOI--DMLRRRAHKPSYHIPQENTGV 387
Db 348 -----LEK-QLRDMADARYQRHMQTVVNSQMKRVTSPPVDHNEQEPSSVA 390
QY 388 EDYQOETRUGSPSPQMPAEVLRKFEAQYKLVKVEKOL-----GLRPSASBPYNO--- 438
Db 391 EMFVDTCTIA---SQSPLLHIP-----KLGISPLQTLSCPGTTLTIQPIILKRPPL 439
QY 439 QBLRNGEPEPRFQELPPFRKNEMKEQYWKQLEIRQOYLN---DMKEIRKMG--REPE 493
Db 440 RKESSLRALHTTATLMKYRAN-----ASSLRTTIDKDGQITSLQKNGTSNOVA 489
QY 494 DIEKDLKQMLONTKESKYPEQ---KYKAKGV-----KFEINLDKCIDENILQEEEA 544
Db 490 DCMNKLHLSLQKKLS--PSELCNRPSPGEGLPNKKVSLVVE-----SDETAVSASG 542
QY 545 MDIPNETLTFEDGMKPKYECVKEHGDYTDKAFELKHCPEAAFTLTWLSFLFLEYSLPH 604
Db 543 ESVPTDSTLTDTKSKSVFVHPSPQSLYVEK-LEKLNIRSDVSKPSKSKTLHGVALPS 601
QY 605 FLEKSPFRSHLIEDLLCANDCSLKDQSEKEMEL 638
Db 602 L---ASPYDVHAEKTIARENE---MDGNFKTWKI 629

RESULT 9
JC7838
Nek6 protein kinase, NIMA histone H3 kinase homolog - human
C:Species: Homo sapiens (man)
C:Date: 03-Dec-2002 #sequence_revision 09-Dec-2002 #text_change 09-Jul-2004
C:Accession: JC7838
R:Hashimoco, Y.; Akita, H.; Hibino, M.; Kohri, K.; Nakanishi, M.
Biochem. Biophys. Res. Commun. 293, 753-758, 2002
A:Title: Identification and characterization of Nek6 protein kinase, a potential human h
A:Reference number: JC7838; MUID:22050088; PMID:12054534
A:Molecule type: mRNA
A:Accession: JC7838
A:Residues: 1-338 <HAS>
A:Cross-references: UNIPROT:O9HC98; DBJ:AB026289
C:Comment: This protein, a novel member of NIMA (never in mitosis, gene A) related kinas
Playing a conserved central role in regulating mitotic chromosome condensation and asse
C:Genetics:
A:Gene: nek6
A:Map position: 9q33-34

Query Match 13.9%; Score 471.5; DB 2; Length 338;
Best Local Similarity 38.5%; Pred. No. 4.3e-12;
Matches 99; Conservative 53; Mismatches 96; Indels 9; Gaps 6;

QY 4 YDVIKAGOGAFKAYLAKGSDSKHCVIKEIN-PEKMPIQBEASKKEVILEKMKHPN 62
Db 70 FOIEKKIGRGQSEVYKATCLLDRTKTVALKVQIFEMDAKARQDCVKEIGLKLQNLHN 129
QY 63 IVAFNFSFOENGLFTVMEYCDGDLMKRIN--RQRGVLFSEDTLGMFVQISLGLKIH 120
Db 130 ILYLDSFIEDNELNLVLELAGDLSQMIKYFKQKRLIPERTVMKVYFVQCSAVEHWH 189
QY 121 DRKILHRDIKAQNFILSKNGVAKLGDGFIARVLNNSMELARTCTGTPYLLSPICQNK 180
Db 190 SREVHRDIKPAVNFITATGVV-KLGDGLGRFFSETTAASHLVCTPYNMSPIRHENG 248
QY 181 YNNKTDIWSLGCVLVELCTLKHPFEGN--NLQOLVLKICQAHFAPISPG--FSRELHSLI 236
Db 249 YNFKSDIWSLGCVLLEYMAALQSPFYGDKNVLFSLCKIEQDCQDYPPL-PGEHYSEKLRLV 307

QY 237 SOLFQVSPDRDRPSINSI 253
Db 308 SMCICPDHPQRPDIGYV 324

RESULT 10
AS7177
NIMA-like protein kinase - Neurospora crassa
C:Species: Neurospora crassa
C:Date: 10-Nov-1995 #sequence_revision 10-Nov-1995 #text_change 11-Aug-2003
C:Accession: AS7177
R:Pu, R.T.; Xu, G.; Wu, L.; Vierula, J.; O'Donnell, K.; Ye, X.S.; Osmani, S.A.
J. Biol. Chem. 270, 18110-18116, 1995
A:Title: Isolation of a functional homolog of the cell cycle-specific NIMA protein kinas
A:Reference number: AS7177; MUID:95355415; PMID:7629122
A:Accession: AS7177
A:Status: preliminary; not compared with conceptual translation
A:Molecule type: DNA
A:Residues: 1-779 <PUA>
A:Cross-references: GB:I42573; NID:G1040682; PIDN:AAA80145.1; PID:G1040683
C:Superfamily: unassigned Ser/Thr or Tyr-specific protein kinases; protein kinase homolo
C:Keywords: ATP
F:5-290/Domain: protein kinase homology <KIN>
F:13-21/Region: protein kinase ATP-binding motif

Query Match 13.5%; Score 459; DB 2; Length 779;
Best Local Similarity 25.2%; Pred. No. 2.7e-11;
Matches 160; Conservative 92; Mismatches 194; Indels 190; Gaps 23;

QY 2 DKYDVIKAGOGAFKAYLAKGSDSKHCVIKEINPEKMPIQBEASKKEVILEKMKHP 61
Db 5 DKYELLEKIGHGSPGIIRKVRKADGMILCRKEITSYLMKSKQEREQLHAEFSILSTLRHP 64
QY 62 NIVAFPN--SFQENGRLFIVMEYCDGDLMKRI-NRQGVLFSEDTLGM--FVQISLGL 116
Db 65 NIVGYHREHLKATQDLHLYMEYCGDGLGRVIRNLKNNQYAESFV-WSIFSQLVATL 123
QY 117 KHIF-----DRKILHRDIKAQNFILSKNGVAKLGDGFI 150
Db 124 YRCHGVDPPEVGVKTLGLGSTARPKPPSGWTILHRDLKPENVFLGEDNSV-KLGDGL 182
QY 151 ARVLNNSMELARTCTGTPYLLSPICONKPYNNKTDIWSLGCVLVELCTLKHPFEGNQL 210
Db 183 SKVM-QSHDFASTYVGTFFYMSPEICAAEKYTLKSDIWSLGCITIELCAREFPFNKATHY 241
QY 211 QVLKICQAHFAPISPGFSRELHSLISOLFQVSPDRDRPSINSILKRPFLNLIPIKYLTPE 270
Db 242 QLVQKIKEGKIAPLPSVYSGELFATIKDLRVNPDPRDPTATLNLV-----IVRLMRKE 296
QY 271 VIOBEFSHMLICRAGAPASRHAGKVQKIOKVRFCGKCPSPRSRISVPIKRNAILHRNE 330
Db 297 KEVVEFSRTL-----RTKEETLNK-RIRELDSKLSALETEKSSIRAEIDASL-RRE 345
QY 331 WRPPAGAKARSIKMIERPKIAAVCGHYDYVYAOQLDMLRRRAHKPSYHIPQENTGVEDY 390
Db 346 WE-----VKAR-----LEIDRL-----VAQE----- 361
QY 391 QGETRHGSPSQMPAEYLQKFEAQYKLVKVEKOL---GLRP-----SSA 432
Db 362 -----IESLQOKFE-QEVOARVEAELORHGRGPMFNSHGQSQGSFSSSTAATL 406
QY 433 EPNTNQRELRSNGEERPFQELPPFRKNEMKEQYWKQLEIRQOYLNLMKEIRKMKR-- 490
Db 407 VSDYNLSVSGDGFPTTDD-----ITDISIAESTDGSIDITKIPRT 450
QY 491 -----EPEDIEK-----DLKQMLQNTK-ESKNPEQYKAKK 521
Db 451 FHRAQTYSSAPAESVLGTPMWDIEMASPSPIIASLSLSPRMALTKAPTINPRIF- 506
QY 522 GVKFEINLDKCIDENILQEEAMDIPEFLTFEDG 557
Db 507 -----GSEPTSTDKNSWNEVPRETEMIDSG 530

submitted to the EMBL Data Library, November 1993

A:Description: Sequencing of chromosome I of Saccharomyces cerevisiae: Analysis of the 42 kbp

A:Reference number: S40891

A:Accession: S40906

A:Molecule type: DNA

A:Residues: 1-430 <CLA>

A:Cross-references: EMBL:L22015; MIPS:YAR018C

R:Clark, M.W.; Keng, T.; Scornis, R.K.; Zhong, W.; Fortin, N.; Zeng, B.; Delaney, S.; Oude

Yeast 10, 535-541, 1994

A:Title: Sequencing of chromosome I of Saccharomyces cerevisiae: analysis of the 42 kbp

A:Reference number: S43441; MUID:95028152; PMID:7941740

A:Accession: S43451

A:Status: nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA

A:Residues: 1-435 <CLW>

A:Cross-references: EMBL:L22015; NID:g1339990; PIDN:AAC04964.1; PID:G349756

A:Note: the nucleotide sequence was submitted to the EMBL Data Library, November 1993

R:Jones, D.G.L.; Rosamond, J.

Gene 90, 87-92, 1990

A:Title: Isolation of a novel protein kinase-encoding gene from yeast by oligodeoxynucleotide

A:Reference number: S11185; MUID:90337351; PMID:2199332

A:Accession: S11185

A:Status: not compared with conceptual translation

A:Molecule type: DNA

A:Residues: 1-97, 'L', 99-120, 'Q', 122-129, 'I', 131-234, 'SQIS', 239-357, 'R', 359-369 <JON>

A:Cross-references: EMBL:MS5416

A:Experimental source: strain SB303

C:Genetics:

A:Gene: SGD:KIN3; NPK1; FUN52

A:Cross-references: SGD:S0000071; MIPS:YAR018C

A:Map position: 1R

C:Superfamily: protein kinase homology

C:Keywords: ATP; phosphotransferase; serine/threonine-specific protein kinase

F:23-343/Domain: protein kinase homology <KIN>

F:31-39/Region: protein kinase ATP-binding motif

Query Match 12.4%; Score 423; DB 2; Length 435;

Best Local Similarity 27.4%; Pred. No. 4.1e-10;

Matches 123; Conservative 69; Mismatches 141; Indels 116; Gaps 14;

QY 3 KYDVKAIGGAGFKAYLAGKSDSKHCYKIKENPKMPIQEKXAEKKEVILKKHFN 62

DB 24 EYVLSEIRGSGSVRKVIHPTKLLVRKDIKYGHMNSKQQIUAECSLSLQKHEN 83

QY 63 IVAFFN-SFOENGR-LPIVMEYCGDGLMKRIN--RQRGVLFSDQILGMFVQISLGLXH 118

DB 84 IVEFYWDPDEQEVLYLYMEYSGDLSQMIKHVQEHKYIPEKIVWGLAQLLTALYK 143

QY 119 IH-----DRK-----ILHRDIKAQNIPLS----- 137

DB 144 CHYGVELPTLTITYDRMKPPVKGNIVIRDLKPGNIFLSYDDSDYININEQVDGHEVNS 203

QY 138 -----KNG-----MVALKGDFGIARVLANSMLARTCTGTPTVYLSPEICQNK 179

DB 204 NYVRDRVNSGKSGSPMDYSQVVVKLGDFGLAKSLSTISQFATTVYGTPTVYMSPEVLMQ 263

QY 180 PYNKTTDINSLGCVLYELCTLKHFPFGNNLQVLVLCQAHFAPISPGFSRELHLSIQL 239

DB 264 PYSPLSDINSLGCVIFEMCSLHPFQAKNYLEQTIKNGKCDTPEYYSRGLNAIISM 323

QY 240 FQVSPDRDSINILKRPFLENLPLKYLPEVIOEEFHSMLICRAGAPASRHKGV--- 296

DB 324 IDVNLATPS-----TFELLQD--IQIRTKSLQLERFERKLLDYE 363

QY 297 -QKCKTQKVRFRGKCPPRGRISVPKRNAILHNEWR-----PPAGAKKARSIKMIERP 349

DB 364 NELTNIEKI-----LEKQALEYEREJLSQLKEQFTQAVEERAREV--ISGK 406

QY 350 KTAAVCGHYDYIYAQLDMLRRRAHKPSYH 378

DB 407 KVGKVPESINGYYG-----KKFAKPAYH 429

RESULT 14

B96587
hypothetical protein F20D21.33 [imported] - Arabidopsis thaliana
C/Species: Arabidopsis thaliana [mouse-ear cross]
C/Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 09-Jul-2004
A/Accession: B96587
R/Theologis, A.; Ecker, J.R.; Palm, C.J.; Pedersen, N.A.; Kaul, S.; White, O.; Alonso,
Chen, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Creasy, T.H.; Dewar, K.;
ansen, N.F.; Hughes, B.; Huizar, L.
Nature 408, 816-820, 2000
A/Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.
C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Lueros, J.S.; Maiti, R.; Marziali,
Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.
A/Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon,
ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.
A/Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.
A/Reference number: A86141; PMID:21016719; PMID:11130712
A/Accession: B96587
A>Status: preliminary
A:Molecule type: DNA
A/Residues: 1-200 <STO>
A/Cross-references: UNIPROT:Q9SLI2; GB:AEO051173; NID:G4585993; PIDN:AAD25629.1; GSPDB:GN
C/Genetics:
A/Gene: F20D21.33
A/Map position: 1

Query Match 12.2% ; Score 415.5 ; DB 2; Length 200;
Best Local Similarity 40.5%; Pred. NO. 4.1e-10; Matches 79; Conservative 51; Mismatches 62; Indels 3; Gaps 3;

Qy 1 MDYDVIKATGAQGKGYAKLKGSDSKHCIVKEINFERMPIQEKAASKVEILLERMKH 60
 ::: |:: |:: |:: |:: |:: |:: |:: |:: |:: |:: |:: |:: |:: |:: |:: |::
Db 1 MEQYEFLQIGKGSFGSALLVRHKHEKKVYLKRARQTQTRFSAHQEVLISKMRH 60
 |:: |:: |:: |:: |:: |:: |:: |:: |:: |:: |:: |:: |:: |:: |:: |::
Qy 61 PNIVAFNSFOENG-RLFIMVECDGDLMKRNQRGVLPSEDQLGHFVQISGLKH 119
 ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||:
Db 61 PFIVEYKDWSWEACVCVICIGVCEGGDAQAKKSNGVHFQEEKLCKWLQJMLGLEYL 120
 ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||:
Qy 120 HDRKIILHRDIKAONIIFLSXNGWVKLGDFGIARLVNNMELARTCTGTVPVLSPEICQNK 179
 ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||:
Db 121 KSNHLIRDRVKNISFLTKEQDI-RLGDFTGLAKIL-TSDDLTSSVVGTSYMCPELLADI 178
 ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||:
Qy 180 PYNNKYTDIWSLGCVL 194
 ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||:
Db 179 PYGKSGDIWSLTFTL 193
 ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||:

RESULT 15

T47988
serine/threonine-protein kinase-like protein - Arabidopsis thaliana
N/Alternate names: protein F21F14.130
C/Species: Arabidopsis thaliana [mouse-ear cross]
C/Date: 20-Apr-2000 #sequence_revision 20-Apr-2000 #text_change 09-Jul-2004
A/Accession: T47988
R/Choise, N.; Robert, C.; Brotter, P.; Wincker, P.; Cattolico, L.; Artiguenave, F.; Sa
submitted to the Protein Sequence Database, February 2000
A/Reference number: Z24481
A/Accession: T47988
A>Status: preliminary
A:Molecule type: DNA
A/Residues: 1-648 <CHO>
A/Cross-references: UNIPROT:Q9M269; EMBL:AL138642
A/Experimental source: cultivar Columbia; BAC clone F21F14
C/Genetics:
A/Map position: 3
A/Introns: 77/3; 117/1; 136/3; 158/2; 189/3; 216/3; 254/1; 272/2; 483/3; 563/1; 586/3
A>Note: F21F14.130

Query Match 12.0% ; Score 410; DB 2; Length 648;
Best Local Similarity 34.2%; Pred. NO. 1.8e-09; Matches 94; Conservative 56; Mismatches 11; Indels 14; Gaps 7;

Qv 10 IGOAGFKAYLAGKSDSKHCIVKEINFERMPIQEKAASKVEILLERMKHPNIVAFFNS 69

GenCore version 5.1.6
Copyright (C) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: November 12, 2004, 12:45:34 ; Search time 198 Seconds
(without alignments)
1859.737 Million cell updates/sec

Title: US-10-730-010-2

Perfect score: 3404

Sequence: 1 MDKDYVKAIGQAGFKAYL.....LCANDCSLKDWSEKEMELRT 640

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1825181 seqs, 575374646 residues

Total number of hits satisfying chosen parameters: 1825181

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0% --

Maximum Match 100%

Listing first 45 summaries

Database : Uniprot 02:*

1: uniprot_sprot:*

2: uniprot_trembl:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB	ID	Description
1	3069	90.2	708	2	Q6P3R8	Q6P3R8 homo sapien
2	3069	90.2	708	2	AAG33885	AAG33885 homo sapi
3	1640.5	48.2	614	2	Q8CCJ0	Q8CCJ0 mus musculus
4	1640.5	48.2	627	2	Q7TSC3	Q7TSC3 mus musculus
5	1135.5	33.4	336	2	Q8C6N6	Q8C6N6 mus musculus
6	1078	31.7	774	1	NEK1_MOUSE	P51954 mus musculus
7	1042	30.6	1258	1	NEK1_HUMAN	Q96PY6 homo sapien
8	1004.5	29.5	424	2	Q8CD72	Q8CD72 mus musculus
9	966	28.4	302	2	Q8BSB6	Q8BSB6 mus musculus
10	850	25.0	898	2	Q7Q028	Q7Q028 giardia lam
11	810.5	23.8	506	1	NEK3_HUMAN	P51956 homo sapien
12	805.5	23.7	511	1	NEK3_MOUSE	Q9R0A5 mus musculus
13	802.5	23.6	509	2	Q99K72	Q99K72 mus musculus
14	802	23.6	510	2	Q6ZN64	Q6ZN64 homo sapien
15	802	23.6	510	2	BAD18511	BAD18511 homo sapi
16	799	23.5	489	2	Q8U023	Q8U023 homo sapien
17	757	22.2	492	2	Q6INC0	Q6INC0 xenopus lae
18	757	22.2	492	2	AAH72363	AAH72363 xenopus l
19	741	21.8	494	2	Q6DJQ5	Q6DJQ5 xenopus tro
20	675.5	19.8	1123	2	Q66T06	Q66T06 dictyosteli
21	675	19.8	781	2	Q6P576	Q6P576 homo sapien
22	675	19.8	781	2	AAH33044	AAH33044 homo sapi
23	675	19.8	849	2	Q7T299	Q7T299 brachydanio
24	668.5	19.6	790	2	Q6NU24	Q6NU24 xenopus lae
25	668.5	19.6	790	2	AAH68778	AAH68778 xenopus l
26	664.5	19.5	911	2	Q6UPR3	Q6UPR3 chlamydomon
27	664.5	19.5	911	2	AAQ64684	AAQ64684 chlamydom
28	659.5	19.4	792	1	NEK4_MOUSE	Q9Z1J2 mus musculus
29	655	19.2	841	1	NEK4_HUMAN	P51957 homo sapien
30	651.5	19.1	792	2	AAH57939	AAH57939 mus muscu
31	638	18.7	692	1	NEK9_HUMAN	Q86896 homo sapien

32	638	18.7	697	2	Q6DJAR	Q6DJAR xenopus tro
33	634	18.6	698	1	NEK8_MOUSE	Q91ZRA mus musculus
34	634	18.6	698	2	AAH70457	AAH70457 mus muscu
35	624	18.3	555	2	Q9N9C3	Q9N9C3 leishmania
36	623	18.3	697	1	NEK8_BRARE	Q90XC2 brachydanio
37	622	18.3	812	2	Q6UPR1	Q6UPR1 chlamydomon
38	622	18.3	812	2	AAQ64686	AAQ64686 chlamydom
39	590	17.3	645	2	Q8NG66	Q8NG66 homo sapien
40	588.5	17.3	637	2	Q8WNU8	Q8WNU8 macaca fasc
41	586.5	17.2	628	2	Q8C0Q4	Q8C0Q4 mus musculus
42	584	17.2	482	2	Q8TBY1	Q8TBY1 homo sapien
43	582.5	17.1	470	2	Q8NG65	Q8NG65 homo sapien
44	575	16.9	579	2	Q95XQ3	Q95XQ3 caenorhabdi
45	571	16.8	365	2	Q8BW62	Q8BW62 mus musculus

ALIGNMENTS

RESULT 1
Q6P3R8
ID Q6P3R8 PRELIMINARY; PRT; 708 AA.
AC Q6P3R8;
DT 05-JUL-2004 (Tremblrel. 27, Created)
DT 05-JUL-2004 (Tremblrel. 27, Last sequence update)
DT 05-JUL-2004 (Tremblrel. 27, Last annotation update)
DE MGC75495 protein.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_taxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Testis;
RX MEDLINE=22388257; PubMed=12477932;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altshul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Haieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Ustin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellaro N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Haie S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
FAhey J., Helton E., Kettaman M., Madan A., Rodriguez S., Sanchez A.,
Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
RA Krzywinski N.I., Skalska U., Smalls D.E., Schnerch A., Schein J.E.,
Jones S.J., Marra M.A.;
RA "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences.";
Proc. Natl. Acad. Sci. U.S.A. 99:16999-16903(2002).
(2)
RP SEQUENCE FROM N.A.
RC TISSUE=Testis;
RA Strausberg R.;
RL Submitted (DEC-2003) to the EMBL/GenBank/DBJ databases.
CC -!- SIMILARITY: Belongs to the Ser/Thr protein kinase family.
EMBL; BC063885; AAH63885.1; -
DR InterPro; IPR011009; Kinase like.
DR InterPro; IPR000719; Prot.kinase.
DR InterPro; IPR002290; Ser_Thr_kinase..
DR InterPro; IPR008271; Ser_Thr_kinase..
DR InterPro; IPR001245; Tyr_kinase.
DR Pfam; PF00069; Pkinase; 1.
DR PRINTS; PR00109; TYRKINASE.
DR ProDom; PD000001; Prot.kinase; 1.
DR SMART; SM00220; S_TKc; 1.
DR SMART; SM00219; TyrKc; 1.
DR PROSITE; PS00107; PROTEIN_KINASE_ATP; UNKNOWN_1.

DR PROSITE; PS50011; PROTEIN KINASE_DOM; 1.
 DR PROSITE; PS50108; PROTEIN KINASE_ST; 1.
 KW ATP-binding; Kinase; Serine/threonine-protein kinase; Transferase.
 SQ SEQUENCE 708 AA; 81445 MW; D9A5B3606414D128 CRC64;

Query Match 90.2%; Score 3069; DB 2; Length 708;
 Best Local Similarity 90.4%; Pred. No. 3.3e-160;
 Matches 585; Conservative 2; Mismatches 4; Indels 56; Gaps 1;

QY 1 MDKYDVKAIGOGAFKAYLAKGSDSKHCVKIENFEKMPIQKEASKKEVILEKMKH 60
 DB 1 MDKYDVKAIGOGAFKAYLAKGSDSKHCVKIENFEKMPIQKEASKKEVILEKMKH 60

QY 61 PNIVAFNSFOENGRLFIWMEYCDGDLMKRINRQGVLFSEDOILGNFVQISLGLKHII 120
 DB 61 PNIVAFNSFOENGRLFIWMEYCDGDLMKRINRQGVLFSEDOILGNFVQISLGLKHII 120

QY 121 DRKILHRDIKAQNIPLSKNGMWAKLGDGFIARVLNNSMELARTCIGTPYILSPICQNK 180
 DB 121 DRKILHRDIKAQNIPLSKNGMWAKLGDGFIARVLNNSMELARTCIGTPYILSPICQNK 180

QY 181 YNNKTDIWSLGCVLVELCTLKHFFEGNQLQVLKICQAHFAPISPGFSRELHSLISOLF 240
 DB 181 YNNKTDIWSLGCVLVELCTLKHFFEGNQLQVLKICQAHFAPISPGFSRELHSLISOLF 240

QY 241 QVSPDRPSINSILKRPFLNLI PKYLTPEVIOEFSHMLICRAGAPASRHAGKVQCK 300
 DB 241 QVSPDRPSINSILKRPFLNLI PKYLTPEVIOEFSHMLICRAGAPASRHAGKVQCK 300

QY 301 IQKVFQKCPSPRSISVPIKRNAILHRNEWPPAGAKARSIKVIERPKIAAVCGHYD 360
 DB 301 IQKVFQKCPSPRSISVPIKRNAILHRNEWPPAGAKARSIKVIERPKIAAVCGHYD 360

QY 361 YVQALDMLRRRAKPSYHPIPOENTGVEDYQETRHGSPSQWPAEYLQKFEAQYK 420
 DB 361 YVQALDMLRRRAKPSYHPIPOENTGVEDYQETRHGSPSQWPAEYLQKFEAQYK 420

QY 421 VEKQLGLRPSAEPNPNQRLSNGEPRFQELPFRKNEMKEQYWKQLEIRQQYLND 480
 DB 421 VEKQLGLRPSAEPNPNQRLSNGEPRFQELPFRKNEMKEQYWKQLEIRQQYLND 480

QY 481 MKEIRKMGREPE-----DIEKDLKQMLQNTKESKNPEQYKAKGVKFEINLDKCIDENILQEEA 493
 DB 481 MKEIRKMGREPE-----DIEKDLKQMLQNTKESKNPEQYKAKGVKFEINLDKCIDENILQEEA 493

QY 545 MDIPNETLTFEDGMKFKEVCEKHEGDTDKAFELKHCPEAAFTLT 591
 DB 601 MDIPNETLTFEDGMKFKEVCEKHEGDTDKAFELKHCPEAGFTQT 647

RESULT 2
 AAH63885 PRELIMINARY; PRT; 708 AA.
 AC AAH63885;
 DT 02-MAR-2004 (Tremblrel. 27, Created)
 DT 02-MAR-2004 (Tremblrel. 27, Last sequence update)
 DT 02-MAR-2004 (Tremblrel. 27, Last annotation update)
 DE MGC75495 protein.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Testis;
 RX MEDLINE=22388257; PubMed=12477932;
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
 RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,

Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh P.,
 Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
 Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
 Brownstein M.J., Udell T.B., Toshiyuki S., Carninci P., Prange C.,
 Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
 Bosak S.A., McKernan K.J., Nalek J.A., Gunaratne P.H.,
 Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
 Villalon D.K., Munz D., Sodergren E.J., Lu X., Gibbs R.A.,
 Fahy J., Helton E., Kettman M., Madan A.C., Rodrigues S., Sanchez A.,
 Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
 Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
 Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
 Krzywinski M.I., Skalska U., Smailus D.E., Scherch A., Schein J.E.,
 Jones S.J., Marra M.A.;
 "Generation and initial analysis of more than 15,000 full-length human
 RT and mouse cDNA sequences.";
 Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).
 [2]
 RL SEQUENCE FROM N.A.
 RP TISSUE=Testis;
 RC Strausberg R.;
 RL Submitted (DEC-2003) to the EMBL/GenBank/DBJ databases.
 DR EMBL; BC063885; AAH63885.1; -;
 SQ SEQUENCE 708 AA; 81445 MW; D9A5B3606414D128 CRC64;

Query Match 90.2%; Score 3069; DB 2; Length 708;
 Best Local Similarity 90.4%; Pred. No. 3.3e-160;
 Matches 585; Conservative 2; Mismatches 4; Indels 56; Gaps 1;

QY 1 MDKYDVKAIGOGAFKAYLAKGSDSKHCVKIENFEKMPIQKEASKKEVILEKMKH 60
 DB 1 MDKYDVKAIGOGAFKAYLAKGSDSKHCVKIENFEKMPIQKEASKKEVILEKMKH 60

QY 61 PNIVAFNSFOENGRLFIWMEYCDGDLMKRINRQGVLFSEDOILGNFVQISLGLKHII 120
 DB 61 PNIVAFNSFOENGRLFIWMEYCDGDLMKRINRQGVLFSEDOILGNFVQISLGLKHII 120

QY 121 DRKILHRDIKAQNIPLSKNGMWAKLGDGFIARVLNNSMELARTCIGTPYILSPICQNK 180
 DB 121 DRKILHRDIKAQNIPLSKNGMWAKLGDGFIARVLNNSMELARTCIGTPYILSPICQNK 180

QY 181 YNNKTDIWSLGCVLVELCTLKHFFEGNQLQVLKICQAHFAPISPGFSRELHSLISOLF 240
 DB 181 YNNKTDIWSLGCVLVELCTLKHFFEGNQLQVLKICQAHFAPISPGFSRELHSLISOLF 240

QY 241 QVSPDRPSINSILKRPFLNLI PKYLTPEVIOEFSHMLICRAGAPASRHAGKVQCK 300
 DB 241 QVSPDRPSINSILKRPFLNLI PKYLTPEVIOEFSHMLICRAGAPASRHAGKVQCK 300

QY 301 IQKVFQKCPSPRSISVPIKRNAILHRNEWPPAGAKARSIKVIERPKIAAVCGHYD 360
 DB 301 IQKVFQKCPSPRSISVPIKRNAILHRNEWPPAGAKARSIKVIERPKIAAVCGHYD 360

QY 361 YVQALDMLRRRAKPSYHPIPOENTGVEDYQETRHGSPSQWPAEYLQKFEAQYK 420
 DB 361 YVQALDMLRRRAKPSYHPIPOENTGVEDYQETRHGSPSQWPAEYLQKFEAQYK 420

QY 421 VEKQLGLRPSAEPNPNQRLSNGEPRFQELPFRKNEMKEQYWKQLEIRQQYLND 480
 DB 421 VEKQLGLRPSAEPNPNQRLSNGEPRFQELPFRKNEMKEQYWKQLEIRQQYLND 480

QY 481 MKEIRKMGREPE-----DIEKDLKQMLQNTKESKNPEQYKAKGVKFEINLDKCIDENILQEEA 493
 DB 481 MKEIRKMGREPE-----DIEKDLKQMLQNTKESKNPEQYKAKGVKFEINLDKCIDENILQEEA 493

QY 545 MDIPNETLTFEDGMKFKEVCEKHEGDTDKAFELKHCPEAAFTLT 591
 DB 601 MDIPNETLTFEDGMKFKEVCEKHEGDTDKAFELKHCPEAGFTQT 647

RESULT 3

Q8CCUQ PRELIMINARY; PRT; 614 AA.

AC Q8CCUQ; 23, Created)

DT 01-MAR-2003 (T-EMBLrel. 23, Last sequence update)

DT 01-MAR-2003 (T-EMBLrel. 23, Last sequence update)

DT 01-MAR-2004 (T-EMBLrel. 26, Last annotation update)

DE Mus musculus 10 days neonate cerebellum cDNA, RIKEN full-length

DE enriched library, clone:653041J22 product:weakly similar to protein

DE kinase nek1 (EC 2.7.1.-).

DE Name=BB049667;

OS Mus musculus (Mouse).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

OX NCBI_taxid=10090;

[1]

SEQUENCE FROM N.A.

RC STRAIN=C57BL/6J; TISSUE=Cerebellum;

RX MEDLINE=99279253; PubMed=10349636;

RA Carninci P., Havaashizaki Y.

RT "High-efficiency full-length cDNA cloning.";

RL Meth. Enzymol. 303:19-44 (1999).

[2]

SEQUENCE FROM N.A.

RC STRAIN=C57BL/6J; TISSUE=Cerebellum;

RX MEDLINE=21085650; PubMed=11217851;

RA RIKEN FANTOM Consortium;

RT "Functional annotation of a full-length mouse cDNA collection.";

RL Nature 409:685-690 (2001).

[3]

SEQUENCE FROM N.A.

RC STRAIN=C57BL/6J; TISSUE=Cerebellum;

RA The FANTOM Consortium.

RA "Analysis of the mouse transcriptome based on functional annotation of

RT 60,770 full-length cDNAs.";

RL Nature 420:563-573 (2002).

[4]

SEQUENCE FROM N.A.

RC STRAIN=C57BL/6J; TISSUE=Cerebellum;

RX MEDLINE=20493374; PubMed=11042159;

RA Carninci P., Shibata Y., Hayatsu N., Sugahara Y., Shibata K., Itoh M.,

RA Konno H., Okazaki Y., Muramatsu M., Hayashizaki Y.;

RT "Normalization and subtraction of cap-trapper-selected cDNAs to

RT prepare full-length cDNA libraries for rapid discovery of new genes.";

RL Genome Res. 10:1617-1630 (2000).

[5]

SEQUENCE FROM N.A.

RC STRAIN=C57BL/6J; TISSUE=Cerebellum;

RX MEDLINE=20530913; PubMed=11076861;

RA Shibata K., Itoh M., Aizawa K., Nagaoka S., Sasaki N., Carninci P.,

RA Konno H., Akiyama J., Nishi K., Itsumai T., Tashiro H., Itoh M.,

RA Sumi N., Ishii Y., Nakamura S., Hazama M., Nishine T., Hatada A.,

RA Yamamoto R., Matsumoto H., Sakaguchi S., Ikegami T., Kashiwagi K.,

RA Fujiwaka S., Inoue K., Togawa Y., Izawa M., Ohara E., Watahiki M.,

RA Yoneda Y., Ishikawa T., Ozawa K., Tanaka T., Matsura S., Kawai J.,

RA Okazaki Y., Muramatsu M., Inoue Y., Kira A., Hayashizaki Y.;

RT "RIKEN integrated sequence analysis (RISA) system-384-format

RT sequencing pipeline with 384 multicapillary sequencer.";

RL Genome Res. 10:1757-1771 (2000).

[6]

SEQUENCE FROM N.A.

RC STRAIN=C57BL/6J; TISSUE=Cerebellum;

RA Adachi J., Aizawa K., Akimura T., Arakawa T., Bono H., Carninci P.,

RA Fukuda S., Furuno M., Hanagaki T., Hara A., Hashizume W.,

RA Hayashida K., Hayatsu N., Hiramoto K., Hiraoka T., Hirozane T.,

RA Hori-F., Imotani K., Ishii Y., Itoh M., Kagawa I., Kasukawa T.,

RA Katoh H., Kawai J., Kojima Y., Kondo S., Konno H., Kouda M., Koya S.,

RA Kurihara C., Matsuyama T., Miyazaki A., Murata M., Nakamura M.,

RA Nishi K., Nomura K., Numazaki R., Ohno M., Ohsato N., Okazaki Y.,

RA Saito R., Saitoh H., Sakai C., Sakai K., Sakazume N., Sano H.,

RA Sasaki D., Shibata K., Shingawa A., Shiraki T., Sogabe Y., Tagami M.,


```

Db 461 NR00W-----DAGAPHTLL 474
RESULT 4
Q7TSC3
ID Q7TSC3 PRELIMINARY; PRT; 627 AA.
AC Q7TSC3;
DT 01-OCT-2003 (TREMELrel. 25, Created)
DT 01-OCT-2003 (TREMELrel. 25, Last sequence update)
DT 01-VAR-2004 (TREMELrel. 26, Last annotation update)
DE BB049667 protein.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
OX NCBI_TaxID=10090;
RN [1]_
RP SEQUENCE FROM N.A.
RC TISSUE=Olfactory epithelium;
RX MEDLINE=22388257; PubMed=12477932;
RA Klausner R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Strausberg R.L., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Heieh F.,
RA Diachenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley D.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A., Sanchez A.,
RA Fahy J., Helton E., Kettner M., Madan A., Rodriguez S., Sanchez A.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
RA Krzywinski M.I., Skalska U., Smailus D.E., Schnerch A., Schein J.E.,
RA Jones S.J., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Olfactory epithelium;
RA Strausberg R.;
RL Submitted (JUN-2003) to the EMBL/GenBank/DBJ databases.
CC -!- SIMILARITY: Belongs to the Ser/Thr protein kinase family.
DR EMBL; BC053516; AAH53516.1; -.
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0004674; F:protein serine/threonine kinase activity; IEA.
DR GO; GO:0016740; F:transferase activity; IEA.
DR GO; GO:0006468; P:protein amino acid phosphorylation; IEA.
DR InterPro; IPR000719; Prot.kinase.
DR InterPro; IPR002290; Ser_thr_kinase.
DR InterPro; IPR008271; Ser_thr_pkin_AS.
DR Pfam; PF00069; Pkinase; 1.
DR ProDom; PD000001; Prot.kinase; 1.
DR SMART; SM00220; S_TKc; 1.
DR PROSITE; PS00107; PROTEIN_KINASE_ATP; UNKNOWN_1.
DR PROSITE; PS00011; PROTEIN_KINASE_DOM; 1.
DR PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
KW ATP-binding; Kinase; Serine/threonine-protein kinase; Transferase.
SQ SEQUENCE 627 AA; 71790 MW; AD579B6621ACA2P CRC64;

Query Match 48.2%; Score 1640.5; DB 2; Length 627;
Best Local Similarity 54.8%; Pred. No. 5.8e-82;
Matches 340; Conservative 49; Mismatches 72; Indels 159; Gaps 9;

QY 1 MDKYDVTKAIGGAFGKAYLAGKSDSKCHVKEINFEKMPIOPEKASKEVILLERQKH 60
Db 1 MDNFELIKIIGEGTFGKYLAKDSSESHCVKEISLTK----EKEASKNEVILLARMEH 56
QY 61 PNIVAFFNSFQENGRFLTVMYCDGGDLMKRINQRGVLPFSEDOILGWFVQISGLKHH 120
Db 57 PNIVTFSSFFQENGRFLTVMYCDGGDLMQRIQRQGVWFSDDQLCWFVQISGLKHH 116

```

```

QY 121 DRKILHRDIKAONIPLSKNGWAKLGDGFIARVLNNSMELARTCTGTPVYLSPEICQKNP 180
Db 117 DRKILHRDIKSONIFLSKNGWAKLGDFTATLNDNSMELAQCTAGTTPYLSPEICQKNP 176
QY 181 YNNKTDINSLGCVLYELCTLKHPPFEGNNLQQLVKICQAHFAPISPGFSRELHSLISOLF 240
Db 177 YNNKTDINSLGCVLYELCTLKHPPFSSNNFHLVLKICQGRVAPISPHFSRDLQSLIPOLF 236
QY 241 QVSPDRPSINSILKRPFLNLIKPYLTPVIOEBSFSEHMLICRACAPASRAGKVVQCK 300
Db 237 RVSPQDRPSVTLKRPFLNLIARSLYPEV----- 267
QY 301 IQKVRPRGKCPPRSISVPIKRNAILHRNWRPPAGAKARSIKMIERFKIAAVCGHYDY 360
Db 268 -----CSR----- 271
QY 361 YVAQLDMLRRRAHKPSYHPIPOENTGVEDYGOETHRGSP-----SQWPAEYLORKEFAQQ 416
Db 272 -----IQSHAH-----MENVAI-----GPTACWRVSPWSAAYLORKEFAQQ 307
QY 417 YLKVKEQLGLRPSAEPNPNORQLNSGEBEPFOELPFRKNEMKEOBYWKOLEIRQQ 476
Db 308 YLKVQRQLGLRPSSEVHPNNEGEKLQSHWETKEQLOYRKNKQKQOBYWKOLEIRQQ 367
QY 477 YLNDMKEIRKQGRPEPEDIKDLQKWLQNTKESXNPQKYKAKKGKVFENLDCISDE 536
Db 368 YHNDMKEIKKMGRE-----LKRWVKEISLDCISDE 400
QY 537 NILQEEAMDINETLTEDGKFKYECVKEHGDYTDKAREKLHCP-----AAF 587
Db 401 DTVQNEAVDKLNATLSFEDGKFGHRRCKBEHEDYTDRAFEELCGPEAGFFQDVIAAE 460
QY 588 TELTWLSFLFLYSLPFL 607
Db 461 NR00W-----DAGAPHTLL 474

RESULT 5
Q8C6N6
ID Q8C6N6 PRELIMINARY; PRT; 336 AA.
AC Q8C6N6;
DT 01-VAR-2003 (TREMELrel. 23, Created)
DT 01-VAR-2003 (TREMELrel. 23, Last sequence update)
DT 01-VAR-2004 (TREMELrel. 26, Last annotation update)
DE Mus musculus 2 days pregnant adult female oviduct cDNA, RIKEN full-length enriched library, clone:E2300241i2 product:weakly similar to protein kinase nek1 (EC 2.7.1.-).
DE Name=BB049667;
GN OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Oviduct;
RX MEDLINE=99279253; PubMed=10349636;
RA Carninci P., Hayashizaki Y.;
RT "High-efficiency full-length cDNA cloning.";
RL Meth. Enzymol. 303:19-44 (1999).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Oviduct;
RX MEDLINE=21085660; PubMed=11217851;
RA RIKEN FANTOM Consortium;
RT "Functional annotation of a full-length mouse cDNA collection.";
RL Nature 409:685-690 (2001).
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Oviduct;
RA The FANTOM Consortium;
RA the RIKEN Genome Exploration Research Group Phase I & II Team;
RT "Analysis of the mouse transcriptome based on functional annotation of

```

60,770 full-length cDNAs.";
 RL Nature 420:563-573 (2002).
 [4]
 RN
 RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6J; TISSUE=Oviduct;
 RX MEDLINE=20499374; PubMed=11042159;
 RA Carninci P., Shibata Y., Hayatsu N., Sugahara Y., Shibata K., Itoh M.,
 RA Kono H., Okazaki Y., Muramatsu M., Hayashizaki Y.;
 RA "Normalization and subtraction of cap-trapper-selected cDNAs to
 RT prepare full-length cDNA libraries for rapid discovery of new genes.";
 RL Genome Res. 10:1617-1630 (2000).
 [5]
 RN
 RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6J; TISSUE=Oviduct;
 RX MEDLINE=20530913; PubMed=11076861;
 RA Shibata K., Itoh M., Aizawa K., Katsunai T., Tashiro H., Itoh M.,
 RA Kono H., Akiyama J., Nishi K., Hatanaka T., Hirozane T.,
 RA Sumi N., Ishii Y., Nakamura S., Hazama M., Nishine T., Harada A.,
 RA Yamamoto R., Matsumoto H., Sakaguchi S., Ikegami T., Kashiwagi K.,
 RA Fujiwaka S., Inoue K., Togawa Y., Izawa M., Ohara E., Watahiki M.,
 RA Yoneda Y., Ishikawa T., Ozawa K., Tanaka T., Matsura S., Kawai J.,
 RA Okazaki Y., Muramatsu M., Inoue Y., Kira A., Hayashizaki Y.;
 RA "RIKEN integrated sequence analysis (RISA) system-384-format
 RT sequencing pipeline with 384 multicapillary sequencer.";
 RL Genome Res. 10:1757-1771 (2000).
 [6]
 RN
 RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6J; TISSUE=Oviduct;
 RX MEDLINE=20530913; PubMed=11076861;
 RA Adachi J., Aizawa K., Akimura T., Arakawa T., Bono H., Carninci P.,
 RA Fukuda S., Furuno M., Hanagaki T., Hara A., Hashizume W.,
 RA Hayashida K., Hayatsu N., Hiramoto K., Hiraoka T., Hirozane T.,
 RA Hori F., Imotani K., Ishii Y., Itoh M., Kagawa I., Kasukawa T.,
 RA Katoh H., Kawai J., Kojima Y., Kondo S., Konno H., Kouda M., Koya S.,
 RA Kurihara C., Matsuyama T., Miyazaki A., Murata M., Nakamura M.,
 RA Nishi K., Nomura K., Numazaki R., Ohno M., Ohsato N., Okazaki Y.,
 RA Saito K., Satoh H., Sakai K., Sakai K., Sakazume N., Sano H.,
 RA Sasaki D., Shibata K., Shingawa A., Shiraki T., Sogabe Y., Tagami M.,
 RA Tagawa A., Takahashi F., Takaku-Akai H., Tanaka Y., Tanaka T.,
 RA Tomaru A., Toya T., Yasunishi A., Muramatsu M., Hayashizaki Y.;
 RL Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.
 CC -1- SIMILARITY: Belongs to the Ser/Thr protein kinase family.
 DR EMBL; AK054168; BAC35877.1; -;
 DR HSP; P31751; IZCK.
 DR MGD; MGI:2142824; BB049667.
 DR GO; GO:0005524; F:ATP binding; IEA.
 DR GO; GO:0004573; F:protein serine/threonine kinase activity; IEA.
 DR GO; GO:0004713; F:protein-tyrosine kinase activity; IEA.
 DR GO; GO:0016740; F:transferase activity; IEA.
 DR GO; GO:0006468; P:protein amino acid phosphorylation; IEA.
 DR InterPro; IPR011009; Kinase like.
 DR InterPro; IPR000719; Prot kinase.
 DR InterPro; IPR002290; Ser Thr pkinase.
 DR InterPro; IPR008271; Ser Thr pkin AS.
 DR InterPro; IPR001245; Tyr_pkinase.
 DR Pfam; PF00069; Pkinase; I.
 DR PRINTS; PR00109; TYRKINASE.
 DR ProDom; PD000001; Prot kinase; 1.
 DR SMART; SM00220; S_TK; 1.
 DR PROSITE; PS00107; PROTEIN KINASE ATP; UNKNOWN_1.
 DR PROSITE; PS00011; PROTEIN KINASE DOM; 1.
 DR PROSITE; PS00108; PROTEIN KINASE ST; 1.
 KW ATP-binding; Kinase; Serine/threonine-protein kinase; Transferase.
 SQ SEQUENCE 336 AA; 38757 MW; 65DD377D72E0143B CRC64;
 Query Match 33.4%; Score 1135.5; Bst 2; Length 336;
 Best Local Similarity 77.0%; Pred. No. 1.4e-54;
 Matches 217; Conservative 22; Mismatches 36; Indels 7; Gaps 2;
 QY 1 MDKVDVITKAGQAFGKAVYAKGSKSKCVIKENFKNPQIEAKSKVILLKWKH 60
 Db 1 MDNFHLIKIIGETFGKVIKADKSSSCVHCIVISLTK-----EKEAKNEVILLARMEH 56
 Y 61 PNTVAFNSFQENGRFLFVMEYCDGDLMKRINRQGVLFSDQILGWVFIISGLKHH 120

```
DR SMART; SM00220; S_TKC; 1.
DR PROSITE; PS00107; PROTEIN KINASE ATP; 1.
DR PROSITE; PS00111; PROTEIN KINASE DOM; 1.
DR PROSITE; PS00108; PROTEIN KINASE ST; 1.
KW ATP-binding; Cell cycle; Cell division; Mitosis; Nuclear protein;
KW Phosphorylation; Serine/threonine-protein kinase; Transferase;
KW Tyrosine-protein kinase.
FT DOMAIN 4 258 Protein kinase.
FT NP_BIND 10 18 ATP (By similarity).
FT BINDING 33 33 ATP (By similarity).
FT ACT_SITE 128 128 Proton acceptor (By similarity).
FT MOD_RES 162 162 Phosphothreonine (by autocatalysis) (By
FT similarity).
SQ SEQUENCE 774 AA; 88427 MW; 7ED5881ACDC06FA CRC64;
Query Match 31.7%; Score 1078; DB 1; Length 774;
Best Local Similarity 36.4%; Pred. No. 5e-51;
Matches 260; Conservative 107; Mismatches 171; Indels 176; Gaps 21;
QY 1 MCKYDVTKAICQAGFKAYLAKGSDSKHCVIKENFEKPIQEKAEKKEVILLEXKH 60
Db 1 MEKYVRLOKIGEGFGKAVLVKSTEDGRHYVYIKENISRMDSKBERQESREVAVLAMKH 60
QY 61 PNIVAFNSFOENGRLFTEVMEYDCGGDLMKRINRQGRVLFSEDOILGWFFVQISGLKHH 120
Db 61 PNIVQYKESFEENGLYIVMDYCEGGDLFKRINAKGALFOEDQILDWFOICLALKHVV 120
QY 121 DRKILHRDIKAQNIPLSKNGMVAKLDPFGIARVLNNSMELARICIGTPPYLSPEIQNK 180
Db 121 DRKILHRDIKSNIFLTGDFGV-QLGDFGIARVLNSTVELARTICIGTPPYLSPEIQNK 179
QY 181 YNNKTDWLSGCVLYELCTLKHPEGNLQVLKICQAHFAPISPFSELSLSLSOLF 240
Db 180 YNNKSDIWLSCVLYELCTLKHAEAGNKNVLKILSGSFPFVPSHYDLSLSLSOLF 239
QY 241 QVSPDRPSINSILKRPFLNIPKYITPTEVQIEFSHMLICRAG----- 285
Db 240 KRNPRDRPSVNSILEKGFIAKRIKFLSPQIIAEFCFLTKLSKFGPQPLPKRPSAQGV 299
QY 286 -----APASHAGKVQKCKIOKVRPGKCPERSIS--VPIKRNAILHREWRP 333
Db 300 SSFVPAQKITPAKYGVPLTYKYGKLLKLEKPPPKHQAHQIPVKX---NNSGEERK 356
QY 334 PAG--AQKARSIKMTER----- 348
Db 357 KMSBEAAKRRLEFIEKEKKQDKQIRFLKAEQMKRQEKQKLERINRAREQGRNVLRA 416
QY 349 -----PKIAVCHYDYVYQAOLDMLR-PAH-----KPSYHP--IPQE 393
Db 417 SGEVKASFFGIGAVSPSPSPGQYHYHAIFDQVQRLAEDNEARKWGGIYGRWLPER 476
QY 384 NTG---VEDYQETRHGSPSPQWPAEYLQKFEAQYKLVKE-----KOLGLRPS 431
Db 477 QKGLHVERANQ-----VEEFQKREAMQNKARAEGHVYLAELRLQIRLQ--- 522
QY 432 AEPNYNQRELR-----SNGEPPFOELPPFRKNEMKEQEVYWKOLEIRQOYLN 479
Db 523 ---NFNERQIQKAKLRGENKEADGKQEAAT-BETDNR--LKQWESLKQATNARAALVK 575
QY 480 DMKIRKQWGPEDIEKDLQKMLQNTKESKNP-----EOKYKAKGVKVF--- 525
Db 576 EQLERKKEAVERE--KKWVEHIVARVSKSDVPLPLELLETGSPSKQVQKPVISVTS 633
QY 526 --EINLDCISDENILOEEAMDIPNLT-----FDGMKFKFEYECVKEH 569
Db 634 LKEVGLDGLTD---TQEEEMKSNAGSISKREILRLNENLKAQDEKEKQH 683
RESULT 7
NEKI_HUMAN STANDARD; PRT; 1258 AA.
AC Q96PY6; Q9Y594;
DT 28-FEB-2003 (Rel. 41, Created)
```

```
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DE 05-JUL-2004 (Rel. 44, Last annotation update)
DE Serine/threonine-protein kinase Nek1 (EC 2.7.1.37) (Nima-related
DE protein kinase 1) (NY-REN-55 antigen).
GN Name=NEK1; Synonyms=KIAA1901;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
NCBI_TaxID=9606;
RN [1]
RN [2]
SEQUENCE FROM N.A. (ISOFORM 1).
TISSUE=Brain;
MEDLINE=21456161; PubMed=11572484;
RX Nagase T., Kikuno R., Ohara O.;
RA Scanlan M.J., Gordan J.D., Williamson B., Stockert E., Bander N.H.,
RA Jongeneel C.V., Gure A.O., Jager D., Jager E., Knuth A., Chen Y.-T.,
RA Old L.J.;
RT "Antigens recognized by autologous antibody in patients with renal-
RT cell carcinoma.";
RL Int. J. Cancer 83:456-464(1999).
CC -!- FUNCTION: Phosphorylates serines and threonines, but also appears
CC to possess tyrosine kinase activity. Implicated in the control of
CC meiosis (By similarity).
CC -!- CATALYTIC ACTIVITY: ATP + a protein = ADP + a phosphoprotein.
CC -!- SUBCELLULAR LOCATION: Nuclear (Probable).
CC -!- ALTERNATIVE PRODUCTS:
CC Event=Alternative splicing; Named isoforms=2;
CC Name=1;
CC IsoId=Q96PY6-1; Sequence=Displayed;
CC Name=2;
CC IsoId=Q96PY6-2; Sequence=VSP_004870;
CC Note=No experimental confirmation available;
CC -!- SIMILARITY: Belongs to the Ser/Thr protein kinase family. NIMA
CC subfamily.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AB067488; BAB67794.1; ALT_INIT.
DR EMBL; AF155113; AAD42879.1; -.
DR HSP; Q13153; 1F3M.
DR Genew; HGNC:7744; NEK1.
DR MIM; 604588; -.
DR InterPro; IPR011009; Kinase like.
DR InterPro; IPR000719; Prot_kinase.
DR InterPro; IPR002290; Ser_thr_pkinase.
DR InterPro; IPR008271; Ser_thr_pkin_AS.
DR Pfam; PF00069; Pkinase; 1.
DR ProDom; PD000001; Prot_kinase; 1.
DR SMART; SM00220; S_TKC; 1.
DR PROSITE; PS00107; PROTEIN KINASE ATP; 1.
DR PROSITE; PS00111; PROTEIN KINASE DOM; 1.
DR PROSITE; PS00108; PROTEIN KINASE ST; 1.
KW Alternative splicing; ATP-binding; Cell cycle; Cell division; Mitosis;
KW Nuclear protein; Phosphorylation; Serine/threonine-protein kinase;
KW Transferase; Tyrosine-protein kinase.
FT DOMAIN 4 258 Protein kinase.
FT NP_BIND 10 18 ATP (By similarity).
FT BINDING 33 33 ATP (By similarity).
FT ACT_SITE 128 128 Proton acceptor (By similarity).
```

```

FT  MOD_RES 162 162 Phosphothreonine (by autocatalysis) (By
FT  VARSPLIC 478 521 similarity).
FT  Missing (in isoform 2).
FT  /FTid=VSP 004870.
FT  CONFLICT 1232 1232 G -> E (in Ref. 2).
FT  SEQUENCE 1258 AA; 142828 MW; 339C4BF8F56612530 CRC64;
SQ
Query Match 30.6%; Score 1042; DB 1; Length 1258;
Best Local Similarity 32.0%; Pred. No. 8.1e-49;
Matches 275; Conservative 121; Mismatches 209; Indels 254; Gaps 24;
QY 1 MDKYDVIKAIIGAGAFKAVLAGKSDSKHCVKIENFERKMPICQEKASKEVILEKWKH 60
Db 1 MEKVLQKIGESFGKALVKSATEDGRQYVKEINISRMSSKREESREAVLANWKH 60
QY 61 PNIVAFNFGQENGRFLIVMEYCDGDLMKRINRQGVLFSEDOQLGNFVQISLGLKHH 120
Db 61 PNIVQVRESFEENGLYIVNDYCEGGDLFKRINAQGVLFQEDQILDWVQICLAKVH 120
QY 121 DRKILHRDIKQNIPLSKNGWAKLQDGFGIARVNNMELARTCTGTPYLSPEICQKP 180
Db 121 DRKILHRDIKQNIPLSKNGWAKLQDGFGIARVNNMELARTCTGTPYLSPEICQKP 180
QY 181 YNKTDIWSLGCYVLYELCTLKHPFEGNLIQQLVLCIQAHPAPISPGFSREHLSISOLF 240
Db 180 YNKSIDIWALGCVLYELCTLKHPFEGNLIQQLVLCIQAHPAPISPGFSREHLSISOLF 240
QY 241 QVSPRRPINSILKPPFLNIPKYLTPVTOEFBSHMLICRAGA---PASRHAG---- 293
Db 240 KNPDRPNSVNSILEKGFIAKIEKFLSPQIAAEFCCLATFSKFSQPIPAKRPAQSGNS 299
QY 294 -KVQCKIKQVFRCKCPPRSRISVPI-----KENAILHNEWRPP----- 334
Db 300 ISWPAQKTK-----PAKGIPLAYKKYDKLHEKRLQKHQAHTPPEKRVNT 351
QY 335 -----AGAKVARSIMIER----- 348
Db 352 GEERRKISEAARKRELFIEKKQKQOIISLMKAEQMKQKQERLERINEAREQGWEN 411
QY 349 -----PKTAACVGHYDYVYQALD-MLRRAHK----- 374
Db 412 VLSAGSGSVKAPFGLSGGTIAPFSRSGQVHYHAIFDQMQQORAEANAKWKREIYG 471
QY 375 ---PSYHPIPQNTGVEDYGOETRH----- 400
Db 472 RGLPERGILPGVRPGF-PYGAAGHHFPDADDIRKTLKELKAVSKQANANRQKQLAVER 530
QY 401 SQWPAEYLQKFEAQYKLVK-----KQLGLRPSAPFNQROELR----- 443
Db 531 AKQVEEFLQKREAVQNKARAECHMVYLLARLRQIRLQ-----NFNERQIQKAKLRGEKK 584
QY 444 ---SNGEE-----PRFQELPRKN-----EMKEQEVWKQLEIRQOYL-- 478
Db 585 EANHSEGOEGSEADMRKIKIESLKAHANARAVALKEQLERKRYAREKXKVVWEHLVA 644
QY 479 --NDMKEIKKKGRBEPDIEKDLKQNR-----LONTKESKNPQK--- 516
Db 645 KGKVSDDVSPPLQGHETGSGSPKQQRMSVISVTSALKKEVGVDLSLTDRETSEEMQKTN 704
QY 517 -YKAKGVKFEINLCKISDENILOEEBAMDIPNLTLPEDQMKPEYCEKHEGDIYDK 575
Db 705 AISSKREILRLN-----ENLKQAEDEKGNLSDTFE-----INVHEDAKEHEKSV 753
QY 576 APEKLHCPEAFETLTWLSFLFLEYSLPHFLLEKSPFSRHLTFDILL----- 621
Db 754 SSDRKKW-EAGQGLVPLDELTLDTSF-----STTERHTVGEVILKGPNGSPRANGK 805
QY 622 CANDSLKDWSEKVELRT 640
Db 806 SPTDSVLKILGAEALQLOT 824

```

RESULT 8

```

Q8CD72
ID Q8CD72 PRELIMINARY; PRT; 424 AA.
AC Q8CD72;
DT 01-VAR-2003 (TrEMBLrel. 23, Created)
DT 01-VAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-VAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Mus musculus 13 days embryo male testis cDNA, RIKEN full-length
DE enriched library, clone:6030407P11 product:NIMA (never in mitosis gene
DE a)-related expressed kinase 1, full insert sequence.
GN Name=Neki;
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Testis;
RX MEDLINE=99279253; Pubmed=10349636;
RA Carninci P., Hayashizaki Y.;
RT "High-efficiency full-length cDNA cloning.";
RL Meth. Enzymol. 303:19-44(1999).
RN [2]
SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Testis;
RX MEDLINE=21085660; Pubmed=11217851;
RA RIKEN FANTOM Consortium;
RT "Functional annotation of a full-length mouse cDNA collection.";
RL Nature 409:685-690(2001).
RN [3]
SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Testis;
RX The FANTOM Consortium;
RA the RIKEN Genome Exploration Research Group Phase I & II Team;
RT "Analysis of the mouse transcriptome based on functional annotation of
RT 60,770 full-length cDNAs.";
RL Nature 420:563-573(2002).
RN [4]
SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Testis;
RX MEDLINE=20499374; Pubmed=11042159;
RA Carninci P., Shibata Y., Hayatsu N., Sugahara Y., Shibata K., Itoh M.,
RA Konno H., Okazaki Y., Muramatsu M., Hayashizaki Y.;
RT "Normalization and subtraction of cap-trapper-selected cDNAs to
RT prepare full-length cDNA libraries for rapid discovery of new genes.";
RL Genome Res. 10:1617-1630(2000).
RN [5]
SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Testis;
RX MEDLINE=20530913; Pubmed=11076861;
RA Shibata K., Itoh M., Aizawa K., Nagao S., Sasaki N., Carninci P.,
RA Konno H., Akiyama J., Nishi K., Kitsuunai T., Tashiro H., Itoh M.,
RA Sumi N., Ishii Y., Nakamura S., Hazama M., Nishire T., Harada A.,
RA Yanamoto R., Matsumoto H., Sakaguchi S., Ikegami T., Kashiwagi K.,
RA Fujiwaka S., Inoue K., Togawa Y., Izawa M., Ohara E., Watanabe M.,
RA Yoneda Y., Ishikawa T., Ozawa K., Tanaka T., Matsuura S., Kawai J.,
RA Okazaki Y., Muramatsu M., Inoue Y., Kira A., Hayashizaki Y.;
RT "RIKEN integrated sequence analysis (RISA) system-384-format
RT sequencing pipeline with 384 multipillar sequencer.";
RL Genome Res. 10:1757-1771(2000).
RN [6]
SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Testis;
RA Adachi J., Aizawa K., Akimura T., Arakawa T., Bono H., Carninci P.,
RA Fukuda S., Furuno M., Hanagaki T., Hara A., Hashizume W.,
RA Hayashida K., Hayatsu N., Hiramoto K., Hiraoka T., Hirozane T.,
RA Horii F., Imotani K., Ishii Y., Itoh M., Kagawa I., Kasukawa T.,
RA Kurihara C., Matsuura T., Miyazaki A., Murata M., Nakamura M.,
RA Nishi K., Nomura K., Numazaki R., Ohno M., Ohsato N., Okazaki Y.,
RA Saito R., Saitoh H., Sakai C., Sakai K., Sakazume N., Sano H.,
RA Sasaki D., Shibata K., Shinagawa A., Shiraki T., Sogabe Y., Tagami M.,
RA Tagawa A., Takahashi F., Takaku-Akashira S., Takeda Y., Tanaka T.,
RA Tomaru A., Toya T., Yasunishi A., Muramatsu M., Hayashizaki Y.;

```


OX	NCBI_TaxID=9606;	DR	ProDom; PD000001; Prot_kinase; 1.
RN	[1]	DR	PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
RP	SEQUENCE FROM N.A.	DR	PROSITE; PS00108; PROTEIN_KINASE_DOM; 1.
RC	TISSUE=Placenta;	DR	PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
RX	MEDLINE=2238257; PubMed=12477932; DOI=10.1073/pnas.242603899;	KW	ATP-binding; Cell cycle; Cell division; Mitosis; Nuclear protein;
RA	Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,	KW	Phosphorylation; Serine/threonine-protein kinase; Transferase.
RA	Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,	FT	DOMAIN 4 257
RA	Hopkins R.F., Jordan H., Moore S.I., Wang J., Hsieh F.,	FT	NP_BIND 10 18
RA	Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,	FT	BINDING 33 33
RA	Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,	FT	ACT_SITE 127 127
RA	Brownstein M.J., Usdin T.B., Tohiyuki S., Carninci P., Prange C.,	FT	MCD_RES 161 161
RA	Rana S.S., McQuellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,	FT	CONFLICT 54 54
RA	Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Guadagnoli P.H.,	FT	CONFLICT 77 78
RA	Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,	FT	CONFLICT 187 189
RA	Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,	FT	CONFLICT 197 189
RA	Fahy J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,	FT	SEQUENCE 506 AA; 57704 MW; 4437EB4A41A4777 CRC64;
RA	Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,	FT	
RA	Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,	FT	
RA	Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,	FT	
RA	Butterfield Y.S.N., Krywinski M.I., Skalska U., Smallus D.E.,	FT	
RA	Schmerch A., Schein J.B., Jones S.J.M., Marra M.A.,	FT	
RA	"Generation and initial analysis of more than 15,000 full-length human	FT	
RT	and mouse cDNA sequences";	FT	
RL	Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).	FT	
RN	[2]	QY	Query Match 23.8%; Score 810.5; DB 1; Length 506;
RP	SEQUENCE OF 48-506 FROM N.A.	DB	Best Local Similarity 34.2%; Pred. No. 1.5e-36;
RC	TISSUE=Placenta;	DB	Matches 193; Conservative 90; Mismatches 169; Indels 113; Gaps 14;
RX	MEDLINE=94368699; PubMed=7522034;	QY	1 MDKYDVIKAIQGAAGKAYLAKGKSDSKHCYKIKENFEKMPLOEKEASKKEVILEMKH 60
RA	Schultz S.J., Fry A.M., Suetterlin C., Ried T., Nigg E.A.;	DB	1 MDDYMLVRLMIGEGSFGALLVQHESSNQMFANKERLPK-SFSNTQSRKEAVLLAKKH 59
RT	"Cell cycle-dependent expression of Nek2, a novel human protein kinase	QY	61 PNIVAFNSFQENGRLFIIVMEYCDGDLMKRINRQGVLFSEDOILGMFVQISLGLKH 120
RT	related to the NIMA mitotic regulator of Aspergillus nidulans.";	DB	60 PNIVAFKESFEAGHLIIVMEYCDGDLQKIKQKGLPDDMLNWFQMCILGVNHI 119
RL	Cell Growth Differ. 5:625-635 (1994).	QY	121 DKILHRDIKQNTPLSKNGVAKLGDGIARVLNLSMELARTCIGTPYVLSPEICNKP 180
RP	SEQUENCE OF 76-189 FROM N.A.	DB	120 KKRVLHRDIKSNIFLTQNGKV-KLGDGSGARLLSNPMFACTVGTPTVYVPPPEIWNLP 178
RX	MEDLINE=94100173; PubMed=8274451;	QY	181 YNNKTDIWSLGCVLVELCTLKHPEGNLQOLVKIKCOAHFAPISPGFSRELHLSOLF 240
RA	Schultz S.J., Nigg E.A.;	DB	179 YNNKSDIWSLGCILVELCTLKHPTQANSKNNLLKVCOCGISPLPSHYVELQFLVKQMF 238
RT	"Identification of 21 novel human protein kinases, including 3 members	QY	241 QVSPDRPSINSILKRPFLNLIKPLTPVIOEFSHMLICRAGAPASRHAGKVQKCK 300
RT	of a family related to the cell cycle regulator nima of Aspergillus	DB	239 KKNPSHRPSATLLSRGIVARLVQKCLPPEIME-----YCEEVLEBEIK 282
RL	nidulans.";	QY	301 IQKVFRGCKPRSRSEISPIKENAILHNENRPPAGAKARSIMKIERPKIAACVGHYD 360
RL	Cell Growth Differ. 4:821-830 (1993).	DB	283 NSKHTNPKTKNPSRIRALGNEASTVOEEQDKGS-----319
CC	-!- FUNCTION: Kinase that may play a role in mitotic regulation.	QY	361 YYAQLDMLRRRAHKPSYHPIPQENTGVEDYQGETRHGSPS-----QWPAEYLQRKFEAQ 416
CC	-!- CATALYTIC ACTIVITY: ATP + a protein = ADP + a phosphoprotein.	DB	320 -HTDLESINENLVESALRRVNRKGG--NKSVHLRKASSPNLHRRQW-----363
CC	-!- SIMILARITY: Belongs to the Ser/thr protein kinase family. NIMA	QY	417 YKLKVEKQLGLPSSAEFNNYQRELRSN--GEPRP-FQELPFRKNEMKEQYWKQLBEI 473
CC	subfamily.	DB	364 -----EKNV---PNTALTALENASILTSSLTAEDDGGSGVSKYKNTTRKQ--W--LKET 411
CC		QY	474 ROQVLDNME-----IRKKVGRE-----PEDIE 496
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration	DB	412 PDLNLILKNADLSLAFQTYIYRPGSEGLKGPLSETEASDSVDGGHSDVILDPERLE 471
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -	QY	497 KDLKQMLQNTKESKNPEQYKAKK 521
CC	the European Bioinformatics Institute. There are no restrictions on its	DB	472 PGLDEEDTDFFEEEDNDPDWVSELKK 496
CC	use by non-profit institutions as long as its content is in no way	RESULT 12	
CC	modified and this statement is not removed. Usage by and for commercial	NEK3_MOUSE	
CC	entities requires a license agreement (See http://www.isb-sib.ch/announce/	ID	NEK3_MOUSE STANDARD; PRT; 511 AA.
CC	or send an email to license@isb-sib.ch).	AC	O9ROA5; Q920X9;
DR	EMBL; BC019916; AAH19916.2; ..	DT	16-OCT-2001 (Rel. 40, Created)
DR	EMBL; Z29067; CAAB2310.1; ..	DT	16-OCT-2001 (Rel. 40, Last sequence update)
DR	EMBL; Z25434; CAAB0921.1; ..	DT	05-JUL-2004 (Rel. 44, Last annotation update)
DR	PIR; I38224; I38224.	DE	Serine/threonine-protein kinase Nek3 (EC 2.7.1.37) (Nima-related
DR	HSSP; P71584; 106Y.	DE	protein kinase 3).
DR	Genew; HGNC:7746; NEK3.	DE	Name=Nek3;
DR	MIM; 604044; ..	OS	Mus musculus (Mouse).
DR	GO; GO:0005634; C:nucleus; NAS.		
DR	GO; GO:0005524; F:ATP binding; NAS.		
DR	GO; GO:0004674; F:protein serine/threonine kinase activity; NAS.		
DR	GO; GO:0007049; P:cell cycle; NAS.		
DR	GO; GO:0007067; P:mitosis; NAS.		
DR	GO; GO:0006468; P:protein amino acid phosphorylation; NAS.		
DR	InterPro; IPR011009; Kinase-like.		
DR	InterPro; IPR000719; Prot_Kinase.		
DR	InterPro; IPR008271; Ser_thr_kin_AS.		
DR	InterPro; IPR001245; Tyr_kinase.		
DR	Pfam; PF00069; Pkinase; 1.		
DR	PRINTS; PR00109; TYRKINASE.		

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: - November 12, 2004, 13:03:10 ; Search time 141 Seconds
(without alignments)
1605.988 Million cell updates/sec

Title: US-10-730-010-2

Perfect score: 3404

Sequence: 1 MDKYDVIRKAGQAFGKAYL.....LCANDCSLKDNSEKEMELRT 640

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1568699 seqs, 353819137 residues

Total number of hits satisfying chosen parameters: 1568699

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Published Applications AA:*
1: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep.*
2: /cgn2_6/ptodata/1/pubpaa/PCT_NEW_PUB.pep.*
3: /cgn2_6/ptodata/1/pubpaa/US05_NEW_PUB.pep.*
4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pep.*
5: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep.*
6: /cgn2_6/ptodata/1/pubpaa/PCTUS_PUBCOMB.pep.*
7: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep.*
8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep.*
9: /cgn2_6/ptodata/1/pubpaa/US09A_PUBCOMB.pep.*
10: /cgn2_6/ptodata/1/pubpaa/US09B_PUBCOMB.pep.*
11: /cgn2_6/ptodata/1/pubpaa/US09C_PUBCOMB.pep.*
12: /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB.pep.*
13: /cgn2_6/ptodata/1/pubpaa/US10A_PUBCOMB.pep.*
14: /cgn2_6/ptodata/1/pubpaa/US10B_PUBCOMB.pep.*
15: /cgn2_6/ptodata/1/pubpaa/US10C_PUBCOMB.pep.*
16: /cgn2_6/ptodata/1/pubpaa/US10D_PUBCOMB.pep.*
17: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB.pep.*
18: /cgn2_6/ptodata/1/pubpaa/US11_NEW_PUB.pep.*
19: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep.*
20: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	ID	Description
1	3404	100.0	640	US-10-243-735-2
2	3404	100.0	640	US-10-730-010-2
3	3081.5	90.5	683	US-09-940-921B-2
4	3081.5	90.5	889	US-10-618-941-94
5	2945	86.5	654	US-09-940-921B-4
6	2107.5	61.9	403	US-10-114-270-188
7	2102.5	61.8	399	US-10-114-270-186
8	1078	31.7	774	US-10-162-706-5
9	1072	31.5	1286	US-10-618-941-95
10	1062.5	31.2	616	US-10-243-735-4
11	1062.5	31.2	616	US-10-730-010-4
12	1060	31.1	1214	US-09-783-320-4
13	1060	31.1	1214	US-10-757-262-124

14	926.5	27.2	254	10	US-09-898-837A-35	Sequence 35, Appl
15	820.5	24.1	506	15	US-10-114-270-180	Sequence 180, Appl
16	815.5	24.0	345	9	US-09-870-962-6	Sequence 6, Appl
17	812	23.9	425	15	US-10-114-270-184	Sequence 184, Appl
18	812	23.9	489	15	US-10-114-270-182	Sequence 182, Appl
19	810.5	23.8	506	17	US-10-618-941-96	Sequence 96, Appl
20	810.5	23.8	506	17	US-10-757-262-118	Sequence 118, Appl
21	808.5	23.8	506	15	US-10-415-011-7	Sequence 7, Appl
22	731	21.5	459	9	US-09-771-161A-198	Sequence 198, Appl
23	675	19.8	850	15	US-10-425-114-56926	Sequence 56926, A
24	675	19.8	850	15	US-10-425-114-56928	Sequence 56928, A
25	663	19.5	841	14	US-10-198-070-52	Sequence 52, Appl
26	663	19.5	841	16	US-10-408-765A-1685	Sequence 1685, Ap
27	663	19.5	841	16	US-10-620-052A-32	Sequence 32, Appl
28	655	19.2	841	14	US-10-198-070-87	Sequence 87, Appl
29	645	18.9	416	9	US-09-731-231A-6	Sequence 6, Appl
30	645	18.9	416	16	US-10-751-985-6	Sequence 6, Appl
31	643	18.8	256	16	US-10-620-052A-48	Sequence 48, Appl
32	638.5	18.8	713	14	US-10-162-706-2	Sequence 2, Appl
33	638	18.7	260	9	US-09-731-231A-5	Sequence 5, Appl
34	638	18.7	260	16	US-10-751-985-5	Sequence 5, Appl
35	638	18.7	692	9	US-09-992-481-2	Sequence 2, Appl
36	638	18.7	692	14	US-10-434-034-2	Sequence 2, Appl
37	636	18.7	692	9	US-09-910-150-5	Sequence 5, Appl
38	636	18.7	692	15	US-10-377-097-5	Sequence 5, Appl
39	634	18.6	698	15	US-10-074-978A-144	Sequence 144, App
40	632.5	18.6	280	14	US-10-162-706-6	Sequence 6, Appl
41	632.5	18.6	460	14	US-10-162-706-4	Sequence 4, Appl
42	632.5	18.6	460	14	US-10-288-798-23	Sequence 23, Appl
43	632.5	18.6	460	15	US-10-362-893-23	Sequence 23, Appl
44	632	18.6	255	10	US-09-898-837A-36	Sequence 36, Appl
45	627.5	18.4	326	15	US-10-074-978A-16	Sequence 16, Appl

ALIGNMENTS

RESULT 1

US-10-243-735-2
; Sequence 2, Application US/10243735
; Publication No. US20030022341A1
; GENERAL INFORMATION:
; APPLICANT: WEBSTER, Marion et al
; TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC
; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES
; TITLE OF INVENTION: THEREOF
; FILE REFERENCE: CL001212DIV
; CURRENT APPLICATION NUMBER: US/10/243,735
; CURRENT FILING DATE: 2002-09-16
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 640
; TYPE: PRT
; ORGANISM: Human
; US-10-243-735-2

Query Match 100.0%; Score 3404; DB 14; Length 640;
Best Local Similarity 100.0%; Pred No 2 6e-215;
Matches 640; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy	1	MDKYDVIRKAGQAFGKAYLAKGKSDSKHCVCVKEINPEKMPIQEKASKEVILLENKMH	60
Dd	1	MDKYDVIRKAGQAFGKAYLAKGKSDSKHCVCVKEINPEKMPIQEKASKEVILLENKMH	60
Qy	61	PNIVAFNSFOENGRPIFVNEYCDGDLMKRINRQGVLFSEQIILGWFIQISLGLKHH	120
Dd	61	PNIVAFNSFOENGRPIFVNEYCDGDLMKRINRQGVLFSEQIILGWFIQISLGLKHH	120
Qy	121	DRKILHRDIKAQNFILSKNGWAKLGDGFGIARVLNNSMELARTICGTPTPLSPFCQNK	180
Dd	121	DRKILHRDIKAQNFILSKNGWAKLGDGFGIARVLNNSMELARTICGTPTPLSPFCQNK	180

QY 181 YNNKTDIWSLGCVLVELCTLKHPPFEGNNLQOLVLKIQAHPAPISPGFSRELHSLISQLF 240
DB 181 YNNKTDIWSLGCVLVELCTLKHPPFEGNNLQOLVLKIQAHPAPISPGFSRELHSLISQLF 240
QY 241 QVSPDRPSINSILKRPFLNLI PKYLTPEVIQEFSSHMLICRAGAPASRHAGVQKCK 300
DB 241 QVSPDRPSINSILKRPFLNLI PKYLTPEVIQEFSSHMLICRAGAPASRHAGVQKCK 300
QY 301 IQKVRFGKCPPRSISVPIKRNAILHNRNWRPAGAKARSIMKIERPKIAAVCGHYDY 360
DB 301 IQKVRFGKCPPRSISVPIKRNAILHNRNWRPAGAKARSIMKIERPKIAAVCGHYDY 360
QY 361 YYAQLDMLRRRAHKSYPHIPOENTGVEDYGOETRHGSPSQWPAEYLQKFEAQYKLL 420
DB 361 YYAQLDMLRRRAHKSYPHIPOENTGVEDYGOETRHGSPSQWPAEYLQKFEAQYKLL 420
QY 421 VEKQGLRPSAEPNQRQELRNGEBEPFOELPFRKNEMKEQYWKQLEETROQVYND 480
DB 421 VEKQGLRPSAEPNQRQELRNGEBEPFOELPFRKNEMKEQYWKQLEETROQVYND 480
QY 481 MKEIRKMGREPEDIKDLKQMLQNTKESKNPQKYKAKKGKVFENLDCISDENILQ 540
DB 481 MKEIRKMGREPEDIKDLKQMLQNTKESKNPQKYKAKKGKVFENLDCISDENILQ 540
QY 541 EEEAMDIPNETLTTPEDGMKFEYECVKEHGDYTDKAFELHCPBAAPTTLTWSLFLFLEY 600
DB 541 EEEAMDIPNETLTTPEDGMKFEYECVKEHGDYTDKAFELHCPBAAPTTLTWSLFLFLEY 600
QY 601 SLPHFLLEKSPFSRHLLIEDLLCANDCSLKDWSEKEMELRT 640
DB 601 SLPHFLLEKSPFSRHLLIEDLLCANDCSLKDWSEKEMELRT 640

RESULT 2

US-10-730-010-2
; Sequence 2, Application US/10730010
; Publication No. US20040142366A1
; GENERAL INFORMATION:
; APPLICANT: WEBSTER, Marion et al
; TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC
; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES
; TITLE OF INVENTION: THEREOF
; FILE REFERENCE: CLO01212CIPDIV-II
; CURRENT APPLICATION NUMBER: US/10730,010
; PRIOR FILING DATE: 2003-12-09
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 640
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-730-010-2

Query Match 100.0%; Score 3404; DB 16; Length 640;
Best Local Similarity 100.0%; Pred. No. 2.6e-215;
Matches 640; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MDKYDVIKAIQGGAFKAYLAGKSDSKHCVIKEINFEKMPIQEKASKKEVILKMKH 60
DB 1 MDKYDVIKAIQGGAFKAYLAGKSDSKHCVIKEINFEKMPIQEKASKKEVILKMKH 60
QY 61 PNIVAFNFSQENGRLFTVMEYCDGGDLMKRINRQGVLFSEDIQILGWVQVQISLGLKHII 120
DB 61 PNIVAFNFSQENGRLFTVMEYCDGGDLMKRINRQGVLFSEDIQILGWVQVQISLGLKHII 120
QY 121 DRKILHRDIKAQNIPLSKNGWVAKLGDGFIARVLNNSMELARTICIGTPYLSPEICQNP 180
DB 121 DRKILHRDIKAQNIPLSKNGWVAKLGDGFIARVLNNSMELARTICIGTPYLSPEICQNP 180
QY 181 YNNKTDIWSLGCVLVELCTLKHPPFEGNNLQOLVLKIQAHPAPISPGFSRELHSLISQLF 240
DB 181 YNNKTDIWSLGCVLVELCTLKHPPFEGNNLQOLVLKIQAHPAPISPGFSRELHSLISQLF 240

QY 241 QVSPDRPSINSILKRPFLNLI PKYLTPEVIQEFSSHMLICRAGAPASRHAGVQKCK 300
DB 241 QVSPDRPSINSILKRPFLNLI PKYLTPEVIQEFSSHMLICRAGAPASRHAGVQKCK 300
QY 301 IQKVRFGKCPPRSISVPIKRNAILHNRNWRPAGAKARSIMKIERPKIAAVCGHYDY 360
DB 301 IQKVRFGKCPPRSISVPIKRNAILHNRNWRPAGAKARSIMKIERPKIAAVCGHYDY 360
QY 361 YYAQLDMLRRRAHKSYPHIPOENTGVEDYGOETRHGSPSQWPAEYLQKFEAQYKLL 420
DB 361 YYAQLDMLRRRAHKSYPHIPOENTGVEDYGOETRHGSPSQWPAEYLQKFEAQYKLL 420
QY 421 VEKQGLRPSAEPNQRQELRNGEBEPFOELPFRKNEMKEQYWKQLEETROQVYND 480
DB 421 VEKQGLRPSAEPNQRQELRNGEBEPFOELPFRKNEMKEQYWKQLEETROQVYND 480
QY 481 MKEIRKMGREPEDIKDLKQMLQNTKESKNPQKYKAKKGKVFENLDCISDENILQ 540
DB 481 MKEIRKMGREPEDIKDLKQMLQNTKESKNPQKYKAKKGKVFENLDCISDENILQ 540
QY 541 EEEAMDIPNETLTTPEDGMKFEYECVKEHGDYTDKAFELHCPBAAPTTLTWSLFLFLEY 600
DB 541 EEEAMDIPNETLTTPEDGMKFEYECVKEHGDYTDKAFELHCPBAAPTTLTWSLFLFLEY 600
QY 601 SLPHFLLEKSPFSRHLLIEDLLCANDCSLKDWSEKEMELRT 640
DB 601 SLPHFLLEKSPFSRHLLIEDLLCANDCSLKDWSEKEMELRT 640

RESULT 3

US-09-940-921B-2
; Sequence 2, Application US/09940921B
; Patent No. US20020147320A1
; GENERAL INFORMATION:
; APPLICANT: Fiddle, Carl Johan
; APPLICANT: Hilbun, Erin
; APPLICANT: Nepomnichy, Boris
; APPLICANT: Hu, Yi
; TITLE OF INVENTION: No. US20020147320A1el Human Kinase Proteins and Polynucleotides Br
; FILE REFERENCE: LEX-0227-USA
; CURRENT APPLICATION NUMBER: US/09/940,921B
; CURRENT FILING DATE: 2002-05-21
; PRIOR APPLICATION NUMBER: US 60/229,280
; PRIOR FILING DATE: 2000-08-31
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 683
; TYPE: PRT
; ORGANISM: homo sapiens
US-09-940-921B-2

Query Match 90.5%; Score 3081.5; DB 9; Length 683;
Best Local Similarity 94.1%; Pred. No. 4.4e-194;
Matches 585; Conservative 2; Mismatches 4; Indels 31; Gaps 1;

QY 1 MDKYDVIKAIQGGAFKAYLAGKSDSKHCVIKEINFEKMPIQEKASKKEVILKMKH 60
DB 1 MDKYDVIKAIQGGAFKAYLAGKSDSKHCVIKEINFEKMPIQEKASKKEVILKMKH 60
QY 61 PNIVAFNFSQENGRLFTVMEYCDGGDLMKRINRQGVLFSEDIQILGWVQVQISLGLKHII 120
DB 61 PNIVAFNFSQENGRLFTVMEYCDGGDLMKRINRQGVLFSEDIQILGWVQVQISLGLKHII 120
QY 121 DRKILHRDIKAQNIPLSKNGWVAKLGDGFIARVLNNSMELARTICIGTPYLSPEICQNP 180
DB 121 DRKILHRDIKAQNIPLSKNGWVAKLGDGFIARVLNNSMELARTICIGTPYLSPEICQNP 180
QY 181 YNNKTDIWSLGCVLVELCTLKHPPFEGNNLQOLVLKIQAHPAPISPGFSRELHSLISQLF 240
DB 181 YNNKTDIWSLGCVLVELCTLKHPPFEGNNLQOLVLKIQAHPAPISPGFSRELHSLISQLF 240
QY 241 QVSPDRPSINSILKRPFLNLI PKYLTPEVIQEFSSHMLICRAGAPASRHAGVQKCK 300

Db 241 QVSPDRPINSILKRPFLNIPKYLTPDEVIOEFSHMLICRAGAPASRHAGKVQCK 300
QY 301 IQKVRFRGKPPRSRISVPIKRNAILHRNWRPPAGAKARSIMIERPKIAAIVCGHYDY 360
Db 301 IQKVRFRGKPPRSRISVPIKRNAILHRNWRPPAGAKARSIMIERPKIAAIVCGHYDY 360
QY 361 YYAQLDMLRRRAHKPSVHPIPOENTGVEDYGOETRHGSPSPQWPAEYLQKFEAQYK 420
Db 361 YYAQLDMLRRRAHKPSVHPIPOENTGVEDYGOETRHGSPSPQWPAEYLQKFEAQYK 420
QY 421 VEKQLGLRSPSSAEPNPNQRLSNGEPRFOELPRKNEMKEQYWKOLEIRQOYLND 480
Db 421 VEKQLGLRSPSSAEPNPNQRLSNGEPRFOELPRKNEMKEQYWKOLEIRQOYLND 480
QY 481 MKEIRKMGREPE-----DIEKDLKQMLRONTKE 509
Db 481 MKEIRKMGREPEENSISHTKTYLVKSNLPVHQDASEGEAPVQDIEKDLKQMLRONTKE 540
QY 510 SKNPEQKYKAGVKVFEINLDCISDENILQEEBAMDIPNETLTFEDGMKFKEVCVKEH 569
Db 541 SKNPEQKYKAGVKVFEINLDCISDENILQEEBAMDIPNETLTFEDGMKFKEVCVKEH 600
QY 570 GDYTDKAFKHLHCPEAAFTLT 591
Db 601 GDYTDKAFKHLHCPEAGFSTQT 622

RESULT 4
US-10-618-941-94
; Sequence 94, Application US/10618941
; Publication No. US20040197792A1
; GENERAL INFORMATION:
; APPLICANT: WHYTE, DAVID
; APPLICANT: MANNING, GERRARD
; APPLICANT: CAENEPEEL, SEAN
; TITLE OF INVENTION: NOVEL KINASES
; FILE REFERENCE: 034536-0321
; CURRENT APPLICATION NUMBER: US/10/618,941
; CURRENT FILING DATE: 2003-07-15
; PRIOR APPLICATION NUMBER: 60/395,632
; PRIOR FILING DATE: 2002-07-15
; NUMBER OF SEQ ID NOS: 143
; SOFTWARE: Patent in version 3.2
; SEQ ID NO 94
; LENGTH: 889
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-618-941-94

Query Match 90.5%; Score 3081.5; DB 17; Length 889;
Best Local Similarity 94.1%; Pred. No. 68-194;
Matches 585; Conservative 2; Mismatches 4; Indels 31; Gaps 1;
QY 1 MDKYDVITKAGGAFGKAYLAKGSDSKHCVIKFKEINPEKMPIQEKASKEVILLEKMKH 60
Db 58 MDKYDVITKAGGAFGKAYLAKGSDSKHCVIKFKEINPEKMPIQEKASKEVILLEKMKH 117
QY 61 PNIVAFNSFQENGRFLFVMEYCDGDLMKRINRQGVLFSDQILGFWFQVLSGLKHH 120
Db 118 PNIVAFNSFQENGRFLFVMEYCDGDLMKRINRQGVLFSDQILGFWFQVLSGLKHH 177
QY 121 DRKILHRDIAQNIPLSKNGWAKLGDGFIARVLNNSMELARTCTGTPTPLSPICQNK 180
Db 178 DRKILHRDIAQNIPLSKNGWAKLGDGFIARVLNNSMELARTCTGTPTPLSPICQNK 237
QY 181 YNNKTDIWSLGCVLVELCTLKHPFEGNNLQVLKICQAHFAPISPGFSRELHSLISOLF 240
Db 238 YNNKTDIWSLGCVLVELCTLKHPFEGNNLQVLKICQAHFAPISPGFSRELHSLISOLF 297
QY 241 QVSPDRPINSILKRPFLNIPKYLTPDEVIOEFSHMLICRAGAPASRHAGKVQCK 300
Db 298 QVSPDRPINSILKRPFLNIPKYLTPDEVIOEFSHMLICRAGAPASRHAGKVQCK 357

QY 301 IQKVRFRGKPPRSRISVPIKRNAILHRNWRPPAGAKARSIMIERPKIAAIVCGHYDY 360
Db 358 IQKVRFRGKPPRSRISVPIKRNAILHRNWRPPAGAKARSIMIERPKIAAIVCGHYDY 417
QY 361 YYAQLDMLRRRAHKPSVHPIPOENTGVEDYGOETRHGSPSPQWPAEYLQKFEAQYK 420
Db 418 YYAQLDMLRRRAHKPSVHPIPOENTGVEDYGOETRHGSPSPQWPAEYLQKFEAQYK 477
QY 421 VEKQLGLRSPSSAEPNPNQRLSNGEPRFOELPRKNEMKEQYWKOLEIRQOYLND 480
Db 478 VEKQLGLRSPSSAEPNPNQRLSNGEPRFOELPRKNEMKEQYWKOLEIRQOYLND 537
QY 481 MKEIRKMGREPE-----DIEKDLKQMLRONTKE 509
Db 538 MKEIRKMGREPEENSISHTKTYLVKSNLPVHQDASEGEAPVQDIEKDLKQMLRONTKE 597
QY 510 SKNPEQKYKAGVKVFEINLDCISDENILQEEBAMDIPNETLTFEDGMKFKEVCVKEH 569
Db 598 SKNPEQKYKAGVKVFEINLDCISDENILQEEBAMDIPNETLTFEDGMKFKEVCVKEH 657
QY 570 GDYTDKAFKHLHCPEAAFTLT 591
Db 658 GDYTDKAFKHLHCPEAGFSTQT 679

RESULT 5
US-09-940-921B-4
; Sequence 4, Application US/09940921B
; Patent No. US20020147320A1
; GENERAL INFORMATION:
; APPLICANT: Fridde, Carl Johan
; APPLICANT: Hilbun, Brin
; APPLICANT: Nepomichy, Boris
; APPLICANT: Hu, Yi
; TITLE OF INVENTION: No. US20020147320A1 Human Kinase Proteins and Polynucleotides E
; FILE REFERENCE: LEX-0227-USA
; CURRENT APPLICATION NUMBER: US/09/940,921B
; CURRENT FILING DATE: 2002-05-21
; PRIOR APPLICATION NUMBER: US 60/229,280
; PRIOR FILING DATE: 2000-08-31
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 654
; TYPE: PRT
; ORGANISM: homo sapiens
US-09-940-921B-4

Query Match 86.5%; Score 2945; DB 9; Length 654;
Best Local Similarity 94.3%; Pred. No. 3.9e-185;
Matches 564; Conservative 7; Mismatches 15; Indels 12; Gaps 4;
QY 1 MDKYDVITKAGGAFGKAYLAKGSDSKHCVIKFKEINPEKMPIQEKASKEVILLEKMKH 60
Db 1 MDKYDVITKAGGAFGKAYLAKGSDSKHCVIKFKEINPEKMPIQEKASKEVILLEKMKH 60
QY 61 PNIVAFNSFQENGRFLFVMEYCDGDLMKRINRQGVLFSDQILGFWFQVLSGLKHH 120
Db 61 PNIVAFNSFQENGRFLFVMEYCDGDLMKRINRQGVLFSDQILGFWFQVLSGLKHH 120
QY 121 DRKILHRDIAQNIPLSKNGWAKLGDGFIARVLNNSMELARTCTGTPTPLSPICQNK 180
Db 121 DRKILHRDIAQNIPLSKNGWAKLGDGFIARVLNNSMELARTCTGTPTPLSPICQNK 180
QY 181 YNNKTDIWSLGCVLVELCTLKHPFEGNNLQVLKICQAHFAPISPGFSRELHSLISOLF 240
Db 181 YNNKTDIWSLGCVLVELCTLKHPFEGNNLQVLKICQAHFAPISPGFSRELHSLISOLF 240
QY 241 QVSPDRPINSILKRPFLNIPKYLTPDEVIOEFSHMLICRAGAPASRHAGKVQCK 300
Db 241 QVSPDRPINSILKRPFLNIPKYLTPDEVIOEFSHMLICRAGAPASRHAGKVQCK 300

APPLICANT: Taupier Jr., Raymond J.
APPLICANT: Casman, Stacie J.
APPLICANT: Ji, Weizhen
APPLICANT: Anderson, David W.
APPLICANT: Liete, Mario W.
APPLICANT: Rastelli, Luca
APPLICANT: Edinger, Shlomit R.
APPLICANT: Stone, David J.
APPLICANT: MacDougall, John R.
APPLICANT: Rothenberg, Mark E.
TITLE OF INVENTION: NO. US20040030110A1 Proteins and Nucleic Acids Encoding Same
FILE REFERENCE: 21402-322C
CURRENT APPLICATION NUMBER: US/10/114,270
CURRENT FILING DATE: 2002-11-27
PRIOR APPLICATION NUMBER: 60/281,086
PRIOR FILING DATE: 2001-04-03
PRIOR APPLICATION NUMBER: 60/281,136
PRIOR FILING DATE: 2001-04-03
PRIOR APPLICATION NUMBER: 60/281,863
PRIOR FILING DATE: 2001-04-05
PRIOR APPLICATION NUMBER: 60/281,906
PRIOR FILING DATE: 2001-04-05
PRIOR APPLICATION NUMBER: 60/282,020
PRIOR FILING DATE: 2001-04-06
PRIOR APPLICATION NUMBER: 60/282,930
PRIOR FILING DATE: 2001-04-10
PRIOR APPLICATION NUMBER: 60/282,934
PRIOR FILING DATE: 2001-04-10
PRIOR APPLICATION NUMBER: 60/283,512
PRIOR FILING DATE: 2001-04-12
PRIOR APPLICATION NUMBER: 60/283,710
PRIOR FILING DATE: 2001-04-13
PRIOR APPLICATION NUMBER: 60/284,234
PRIOR FILING DATE: 2001-04-17
Remaining Prior Application data removed - See File Wrapper or PALM.
NUMBER OF SEQ ID NOS: 470
SEQ ID NO 186
LENGTH: 399
TYPE: PRT
ORGANISM: Homo sapiens
US-10-114-270-186

Query Match 61.8%; Score 2102.5; DB 15; Length 399;
Best Local Similarity 99.0%; Pred. No. 58-130; 0; Indels 3; Gaps 1;
Matches 398; Conservative 1; Mismatches 0

QY 1 MDKYDVIRKAGGAFKAYLAKGSDSKHCVCVKEINFERKMPIQEKEASKKEVILLERKMH 60
DB 1 MDKYDVIRKAGGAFKAYLAKGSDSKHCVCVKEINFEK---QEKEASKKEVILLERKMH 57

QY 61 PNIVAFNFSQENGRLFIYMEYCDGDLMKRINRQGVLFSEDOILGWFVQISLGLKHH 120
DB 58 PNIVAFNFSQENGRLFIYMEYCDGDLMKRINRQGVLFSEDOILGWFVQISLGLKHH 117

QY 121 DRKILHRDIKAQNIKQVAFKAGKSDSKHCVCVKEINFERKMPIQEKEASKKEVILLERKMH 180
DB 118 DRKILHRDIKAQNIKQVAFKAGKSDSKHCVCVKEINFERKMPIQEKEASKKEVILLERKMH 177

QY 181 YNNKTDIWSLGCVLVELCTLKHPFEGNQLQVLKICQAHFAPISPGFSRELHSLISOLF 240
DB 178 YNNKTDIWSLGCVLVELCTLKHPFEGNQLQVLKICQAHFAPISPGFSRELHSLISOLF 237

QY 241 QVSPDRPSINSILKRPFLNLPKYLTPVIEQESFHMILCRAGAPASRHAGKVQCK 300
DB 238 QVSPDRPSINSILKRPFLNLPKYLTPVIEQESFHMILCRAGAPASRHAGKVQCK 297

QY 301 IQKVRFGKCPSPRSISVPIKRNAILHRNWRPAGAKARSIMIERPKIAAVCGHYDY 360
DB 298 IQKVRFGKCPSPRSISVPIKRNAILHRNWRPAGAKARSIMIERPKIAAVCGHYDY 357

QY 361 YYAQLDMLRRRAKHSYHPIDQENTGVEDYGOETRHGSPSQ 402
DB 358 YYAQLDMLRRRAKHSYHPIDQENTGVEDYGOETRHGSPSQ 399

RESULT 8
US-10-162-706-5
Sequence 5, Application US/10162706
Publication No. US20030059918A1
GENERAL INFORMATION:
APPLICANT: Smolyar, Alex
TITLE OF INVENTION: Regulation of Human Serine/Threonine
TITLE OF INVENTION: Kinase
FILE REFERENCE: 004974.00731
CURRENT APPLICATION NUMBER: US/10/162,706
CURRENT FILING DATE: 2002-06-06
PRIOR APPLICATION NUMBER: US 60/296,164
PRIOR FILING DATE: 2001-06-07
PRIOR APPLICATION NUMBER: US 60/323,100
PRIOR FILING DATE: 2001-09-19
PRIOR APPLICATION NUMBER: US 60/330,578
PRIOR FILING DATE: 2001-10-25
PRIOR APPLICATION NUMBER: US 60/348,601
PRIOR FILING DATE: 2002-01-17
NUMBER OF SEQ ID NOS: 18
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 5
LENGTH: 774
TYPE: PRT
ORGANISM: Homo sapiens
US-10-162-706-5

Query Match 31.7%; Score 1078; DB 14; Length 774;
Best Local Similarity 36.4%; Pred. No. 2.3e-62;
Matches 260; Conservative 107; Mismatches 171; Indels 176; Gaps 21;

QY 1 MDKYDVIRKAGGAFKAYLAKGSDSKHCVCVKEINFERKMPIQEKEASKKEVILLERKMH 60
DB 1 MEKYVRLQKIGEGSFGKAVLVKSTEDGRHYVKEINISRMDSKQESRESEVAVLANWKH 60

QY 61 PNIVAFNFSQENGRLFIYMEYCDGDLMKRINRQGVLFSEDOILGWFVQISLGLKHH 120
DB 61 PNIVQYKESPEENGSLYVMDYCEGDLFRINAKGALFQEDQILDWFVQICLALXHVH 120

QY 121 DRKILHRDIKAQNIKQVAFKAGKSDSKHCVCVKEINFERKMPIQEKEASKKEVILLERKMH 180
DB 121 DRKILHRDIKAQNIKQVAFKAGKSDSKHCVCVKEINFERKMPIQEKEASKKEVILLERKMH 179

QY 181 YNNKTDIWSLGCVLVELCTLKHPFEGNQLQVLKICQAHFAPISPGFSRELHSLISOLF 240
DB 180 YNNKSDIWLGCVLVELCTLKHPFEGNQLQVLKICQAHFAPISPGFSRELHSLISOLF 239

QY 241 QVSPDRPSINSILKRPFLNLPKYLTPVIEQESFHMILCRAGAPASRHAGKVQCK 285
DB 240 KRNPRDRPSINSILKRPFLNLPKYLTPVIEQESFHMILCRAGAPASRHAGKVQCK 299

QY 286 ---APASRHAGKVQCKIQVFRGKCPSPRSIS--VPIKRNAILHRNWRP 333
DB 300 SSFVPAKIKTPAAKYGVPLTYKYGDKLLEKPPPKHQAQHPVK---KNSGEERK 356

QY 334 PAG--AQKARSIMIER----- 348
DB 357 KMSEAAKRLRLEFIEKEKQKQIRFLKAEQMKRQKRLERINRAREQGNVLRAGG 416

QY 349 ---PKIAAVCGHYDYVAQDMLRR--RAH-----KPSYHP--IPQE 383
DB 417 SGEVKASFFGIGGAVSPSPCPRGQYEHYHAIQDQVORLEADNEARWKGGIYGRWLPER 476

QY 384 NTG---VEDYGOETRHGSPSQWPAEYLRQKFAQQYKLVKVE-----KQJGRPSS 431
DB 477 QKGLHAVERANQ-----VEEFLQRKRAMQKARAEHVYVYVLAELRQIRLQ--- 522

QY 432 AEPNYNORQELR-----SNGEPRFQELPFRKMEKQEQVYKQLEIRQOVLN 479
DB 523 ---NFNERQQTAKLGRGENKEADGTGQGEAT--BETDNR---LKMSLSKATQNRARAVLK 575

Db 357 KMSEAAKRRLEFIEKEKKQKQDQIRFLKAEQMKRQEKQRLERINRAREQGWRNVLRAGG 416
Qy 349 -----PKIAAVCGHYDYVYAAQDMLRR-RAH-----KPSVHP--IPQE 383
Db 417 SGVKASFGIGAVSPSPCSFGQYEHYHALFDQMLRAEDNEARWKGGIYGRWLPER 476
Qy 384 NTG---VEDYGQETHRGSPSPQWPAEYLQKPEAQYKLVKVE-----KQLGLRPS 431
Db 477 QKGLAVERANQ-----VEEFLQKREAMQNKARAEHGHHVVLARLRQIRLQ--- 522
Qy 432 ASPNNYORQLRNSGEBRPFQBLPFRKNEMKEQYWKQLEBIRQYLNDMKEIRKKVGRE 491
Db 523 ---NFERQOIKA-----LRGENKEADGTK-----GQE 548
Qy 492 PEDIEKDLKOMRLQNTKESKNP-----EOKYAKKGVKFEINLDKICISDENILOEEBAMD 546
Db 549 ATE-ETDMLKXVESLKQAQTNARAVALKQLEKRRKEAYE--REKKVWEHLVARVKSSD 605
Qy 547 IPNETLTFEDG 557
Db 606 VPLPLELLETG 616

RESULT 11
US-10-730-010-4
; Sequence 4, Application US/10730010
; Publication No. US20040142366A1
; GENERAL INFORMATION:
; APPLICANT: WEBSTER, Marion et al
; TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC
; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USBS
; FILE REFERENCE: CLO01212CIPDIV-II
; CURRENT APPLICATION NUMBER: US/10/730,010
; CURRENT FILING DATE: 2003-12-09
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 616
; TYPE: PRT
; ORGANISM: Mouse
US-10-730-010-4

Query Match 31.2%; Score 1062.5; DB 16; Length 616;
Best Local Similarity 36.7%; Pred. No. 1.8e-61;
Matches 246; Conservative 108; Mismatches 148; Indels 169; Gaps 18;
Qy 1 MDKYDVKAIGQAGAFKAYLAKGSDSKHCVKINEFEMPIQEKASKEVILLKMKH 60
Db 1 MEKYVRLQKIGSGFGKAVLVKSTEDGRHYVIKEINISRMDSKESRREAVLANMKH 60
Qy 61 PNIVAFNSFOENGELFTVMEXCDGDLMKRINRQGVLFSEDOILGHFWQISLGLKH 120
Db 61 PNIVQKSGFEENGSLYIVMDYCEGGDLFKRINAQKALFOEDQILDWFVQICLAKHVH 120
Qy 121 DRKILHRDIKAQNIPLSKNGMVAKLGDGFIARVLNNSMELARTICGTPTYYLSPECNKP 180
Db 121 DRKILHRDIKSNIFLTNDGTV-QLGDGFIARVLNSTVELARTICGTPTYYLSPECNKP 179
Qy 181 YNNKTIDWSLGCVLVELCTLKHFFPGNNLQOLVLKICQAHFAPISPGPSRELHLSOLF 240
Db 180 YNNKSDIHALGCVLYELCTLKHAFAGNKNLVLKIIISGSPFPVPHYSYDLRSLLSOLF 239
Qy 241 QVSPDRDRPSINSILKRPFLNLIPLYLTPVETQEEFSHMLICRAG----- 285
Db 240 XENPRDRPSVNSILKGFIAKRIEFLSPQLIAEFCLTKSLKSGQPLPKRPSAGGV 299
Qy 286 -----APASHAGKWKQCKIQKVRFGKCPSPRSIS--VPIKRNAILHRNEWRP 333
Db 300 SSFVPAQKITKPAKYGYVPLTVKYGDKKLEKKPPKHKQAHQIPVKK---MNSGEERK 356
Qy 334 PAG--AQXARSITMLER----- 348

Db 357 KMSEAAKRRLEFIEKEKKQKQDQIRFLKAEQMKRQEKQRLERINRAREQGWRNVLRAGG 416
Qy 349 -----PKIAAVCGHYDYVYAAQDMLRR-RAH-----KPSVHP--IPQE 383
Db 417 SGVKASFGIGAVSPSPCSFGQYEHYHALFDQMLRAEDNEARWKGGIYGRWLPER 476
Qy 384 NTG---VEDYGQETHRGSPSPQWPAEYLQKPEAQYKLVKVE-----KQLGLRPS 431
Db 477 QKGLAVERANQ-----VEEFLQKREAMQNKARAEHGHHVVLARLRQIRLQ--- 522
Qy 432 ASPNNYORQLRNSGEBRPFQBLPFRKNEMKEQYWKQLEBIRQYLNDMKEIRKKVGRE 491
Db 523 ---NFERQOIKA-----LRGENKEADGTK-----GQE 548
Qy 492 PEDIEKDLKOMRLQNTKESKNP-----EOKYAKKGVKFEINLDKICISDENILOEEBAMD 546
Db 549 ATE-ETDMLKXVESLKQAQTNARAVALKQLEKRRKEAYE--REKKVWEHLVARVKSSD 605
Qy 547 IPNETLTFEDG 557
Db 606 VPLPLELLETG 616

RESULT 12
US-09-783-320-4
; Sequence 4, Application US/09783320
; Patent No. US20020038011A1
; GENERAL INFORMATION:
; APPLICANT: Walke, D. Wade
; APPLICANT: Hu, Yi
; APPLICANT: Nepomnichy, Boris
; APPLICANT: Turner, C. Alexander Jr
; APPLICANT: Zambrowicz, Brian
; TITLE OF INVENTION: No. US20020038011A1 Human Kinases and Polynucleotides Encoding
; FILE REFERENCE: LEX-0137-USA
; CURRENT APPLICATION NUMBER: US/09/783,320
; CURRENT FILING DATE: 2001-02-15
; PRIOR APPLICATION NUMBER: US 60/183,582
; PRIOR FILING DATE: 2000-02-18
; PRIOR APPLICATION NUMBER: US 60/184,014
; PRIOR FILING DATE: 2000-02-22
; NUMBER OF SEQ ID NOS: 50
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 1214
; TYPE: PRT
; ORGANISM: homo sapiens
US-09-783-320-4

Query Match 31.1%; Score 1060; DB 9; Length 1214;
Best Local Similarity 33.0%; Pred. No. 5.8e-61;
Matches 270; Conservative 128; Mismatches 203; Indels 218; Gaps 22;
Qy 1 MDKYDVKAIGQAGAFKAYLAKGSDSKHCVKINEFEMPIQEKASKEVILLKMKH 60
Db 1 MEKYVRLQKIGSGFGKAVLVKSTEDGRQYVIKEINISRMSSKESRREAVLANMKH 60
Qy 61 PNIVAFNSFOENGELFTVMEXCDGDLMKRINRQGVLFSEDOILGHFWQISLGLKH 120
Db 61 PNIVQKSGFEENGSLYIVMDYCEGGDLFKRINAQKALFOEDQILDWFVQICLAKHVH 120
Qy 121 DRKILHRDIKAQNIPLSKNGMVAKLGDGFIARVLNNSMELARTICGTPTYYLSPECNKP 180
Db 121 DRKILHRDIKSNIFLTNDGTV-QLGDGFIARVLNSTVELARTICGTPTYYLSPECNKP 179
Qy 181 YNNKTIDWSLGCVLVELCTLKHFFPGNNLQOLVLKICQAHFAPISPGPSRELHLSOLF 240
Db 180 YNNKSDIHALGCVLYELCTLKHAFAGNKNLVLKIIISGSPFPVPHYSYDLRSLLSOLF 239
Qy 241 QVSPDRDRPSINSILKRPFLNLIPLYLTPVETQEEFSHMLICRAG---PASHAG--- 293
Db 240 XENPRDRPSVNSILEKGFIAKRIEFLSPQLIAEFCLTKFSKFGSQPIPAKEPASQNS 299

GenCore version 5.1.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 12, 2004, 12:54:29 ; Search time 40 Seconds
(without alignments)
1061.088 Million cell updates/sec

Title: US-10-730-010-2

Perfect score: 3404

Sequence: 1 MDKYDVIRKAIQGGAFKAYL.....LCANDCSLKDWSEKEMELRT 640

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 478139 segs, 66318000 residues

Total number of hits satisfying chosen parameters: 478139

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents AA:*

1: /cgn2_6/prodata/1/aaa/5A-COMB.pep:*

2: /cgn2_6/prodata/1/aaa/5B-COMB.pep:*

3: /cgn2_6/prodata/1/aaa/6A-COMB.pep:*

4: /cgn2_6/prodata/1/aaa/6B-COMB.pep:*

5: /cgn2_6/prodata/1/aaa/PCTUS-COMB.pep:*

6: /cgn2_6/prodata/1/aaa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	3404	100.0	640	4	US-09-873-404-2
2	3404	100.0	640	4	US-10-243-735-2
3	1062.5	31.2	616	4	US-09-873-404-4
4	1062.5	31.2	616	4	US-10-243-735-4
5	815.5	24.0	345	3	US-09-173-581-6
6	815.5	24.0	345	3	US-09-420-915-6
7	731	21.5	459	4	US-09-538-092-1182
8	655	19.2	841	4	US-09-538-092-1183
9	638	18.7	692	4	US-09-992-481-2
10	590	17.3	645	4	US-10-196-927-2
11	594	17.2	482	4	US-10-196-927-4
12	538.5	15.8	979	3	US-08-870-529-2
13	538.5	15.8	979	4	US-09-544-794-2
14	535.5	15.7	445	4	US-09-167-206-2
15	535.5	15.7	445	4	US-09-538-092-1181
16	512.5	15.1	699	3	US-09-457-040B-18
17	462.5	13.6	302	3	US-09-221-235-2
18	462.5	13.6	302	3	US-09-221-928-2
19	462.5	13.6	302	3	US-09-221-527-2
20	462.5	13.6	302	3	US-09-221-236-2
21	462.5	13.6	302	3	US-09-221-416-2
22	462.5	13.6	302	3	US-09-221-245-2
23	462.5	13.6	302	3	US-09-163-115-2
24	462.5	13.6	302	3	US-09-221-528-2
25	462.5	13.6	302	3	US-09-593-553-2
26	462.5	13.6	302	3	US-09-221-237-2
27	462.5	13.6	302	4	US-09-757-982-2

28	458.5	13.5	302	4	US-09-579-664B-9	Sequence 9, Appli
29	458.5	13.5	302	4	US-10-355-975A-9	Sequence 9, Appli
30	450.5	13.2	451	4	US-09-248-796A-18414	Sequence 18414, A
31	427	12.5	233	2	US-08-712-709-1	Sequence 1, Appli
32	427	12.5	233	3	US-09-111-444-1	Sequence 1, Appli
33	427	12.5	233	3	US-09-541-228-1	Sequence 1, Appli
34	415.5	12.2	647	3	US-09-031-563-7	Sequence 7, Appli
35	415.5	12.2	647	4	US-09-392-277-7	Sequence 7, Appli
36	415.5	12.2	647	4	US-09-258-000-7	Sequence 7, Appli
37	415.5	12.2	648	3	US-09-031-563-5	Sequence 5, Appli
38	415.5	12.2	648	4	US-09-392-277-5	Sequence 5, Appli
39	415.5	12.2	648	4	US-09-258-000-5	Sequence 5, Appli
40	415.5	12.2	1115	4	US-10-335-711-2	Sequence 2, Appli
41	415.5	12.2	1315	3	US-09-031-563-2	Sequence 2, Appli
42	415.5	12.2	1315	3	US-09-293-505-10	Sequence 10, Appli
43	415.5	12.2	1315	4	US-09-392-277-2	Sequence 2, Appli
44	415.5	12.2	1315	4	US-09-258-000-2	Sequence 2, Appli
45	415.5	12.2	1315	4	US-09-660-939A-10	Sequence 10, Appli

ALIGNMENTS

RESULT 1

US-09-873-404-2

; Sequence 2, Application US/09873404

; Patent No. 6500656

; GENERAL INFORMATION:

; APPLICANT: WESTER, Marion et al

; TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC

; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES

; TITLE OF INVENTION: THEREOF

; FILE REFERENCE: CUC01212-CIP

; CURRENT APPLICATION NUMBER: US/09/873,404

; CURRENT FILING DATE: 2001-06-05

; NUMBER OF SEQ ID NOS: 4

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 2

; LENGTH: 640

; TYPE: PRT

; ORGANISM: Human

; US-09-873-404-2

Query Match		100.0%;	Score 3404;	DB 4;	Length 640;
Best Local Similarity		100.0%;	Pred. No. 2.3e-272;	Mismatches 0;	Indels 0;
Matches 640;		Conservative 0;			
Qy	1	MDKYDVIRKAIQGGAFKAYLAKGSDSKHCVCIKKEINFEKMPIQEKESKEVILLEKMKH	60		
Db	1	MDKYDVIRKAIQGGAFKAYLAKGSDSKHCVCIKKEINFEKMPIQEKESKEVILLEKMKH	60		
Qy	61	PNIVAFNSFOENGRLFIYMEYCDGDLMKRINRQGVLPSEDOILGWFVQISLGLKHH	120		
Db	61	PNIVAFNSFOENGRLFIYMEYCDGDLMKRINRQGVLPSEDOILGWFVQISLGLKHH	120		
Qy	121	DRKTLHREDIKAQIFLSKNGWAKLGDGFIARVLNNSMELARTCIGTPYVLSPEIQNKP	180		
Db	121	DRKTLHREDIKAQIFLSKNGWAKLGDGFIARVLNNSMELARTCIGTPYVLSPEIQNKP	180		
Qy	181	YNNKTDIWSLGCVLVELCTLKHFPFEGNNLOQLVLIKQAHFAPISFGFSELHSLISQLF	240		
Db	181	YNNKTDIWSLGCVLVELCTLKHFPFEGNNLOQLVLIKQAHFAPISFGFSELHSLISQLF	240		
Qy	241	QVSPDRPSTNSILKRPFLNLI PKYLTPEVIOEFSSHMLICRAGAPASHAGKVKQCK	300		
Db	241	QVSPDRPSTNSILKRPFLNLI PKYLTPEVIOEFSSHMLICRAGAPASHAGKVKQCK	300		
Qy	301	IQVRFKGCPPSRISVPIKRNAILHRNEWRPPAGAKARSIMTERPKIAAVCGHYDY	360		
Db	301	IQVRFKGCPPSRISVPIKRNAILHRNEWRPPAGAKARSIMTERPKIAAVCGHYDY	360		
Qy	361	YYAQLDMLRRRAKPSVHPPIQENTGVEDYQETRHGPSQWPAEYLRQKFEAQYKLLK	420		

Db 361 YYAQLDMLRRRAHPSYHP1PQENTGVEDYQETRHGSPSQWPAEYLRKFAQQYK 420
QY 421 VEKQLGLRPSAEPNYNQRELRSNGBEPRFQELPFRKNEMKEQYWKQLEERQQVYND 480
Db 421 VEKQLGLRPSAEPNYNQRELRSNGBEPRFQELPFRKNEMKEQYWKQLEERQQVYND 480
QY 481 MKEIRKMGREPEDIKDLKQMLQNTKESKNPEQYKAKKGKVFPEINLDCISDENILQ 540
Db 481 MKEIRKMGREPEDIKDLKQMLQNTKESKNPEQYKAKKGKVFPEINLDCISDENILQ 540
QY 541 EEEAMDIPNEILTTFEDGKMKFKEVCVKEHGDYTDKAFKHLCPAAFTETLWLSFLFLEY 600
Db 541 EEEAMDIPNEILTTFEDGKMKFKEVCVKEHGDYTDKAFKHLCPAAFTETLWLSFLFLEY 600
QY 601 SLPHFLLEKSPFSRHLLIEDLLCANDCSLKDWSEKEMELRT 640
Db 601 SLPHFLLEKSPFSRHLLIEDLLCANDCSLKDWSEKEMELRT 640
RESULT 2
US-10-243-735-2
; Sequence 2, Application US/10243735
; Patent No. 6706510
; GENERAL INFORMATION:
; APPLICANT: WEBSTER, Marion et al
; TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC
; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES
; TITLE OF INVENTION: THEREOF
; FILE REFERENCE: CL001212DIV
; CURRENT APPLICATION NUMBER: US/10/243,735
; CURRENT FILING DATE: 2002-09-16
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 640
; TYPE: PRT
; ORGANISM: Human
US-10-243-735-2
Query Match 100.0%; Score 3404; DB 4; Length 640;
Best Local Similarity 100.0%; Pred. No. 2.3e-272;
Matches 640; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MDKYDV1KAIQGGAGFKAYLAKGSDSKHCVIKKEINFEKMPIQEKASKEVILEKMKH 60
Db 1 MDKYDV1KAIQGGAGFKAYLAKGSDSKHCVIKKEINFEKMPIQEKASKEVILEKMKH 60
QY 61 PNIVAFNPSFQENGRLEFVMEYCDGGDLMKRINRQGVLFSEDOILGWFWQISLGLKH 120
Db 61 PNIVAFNPSFQENGRLEFVMEYCDGGDLMKRINRQGVLFSEDOILGWFWQISLGLKH 120
QY 121 DRKILHRDIKAQNTFLSKNGVAKLGFGIARVLNNSMELARTCIGTPPYLSPEICQNK 180
Db 121 DRKILHRDIKAQNTFLSKNGVAKLGFGIARVLNNSMELARTCIGTPPYLSPEICQNK 180
QY 181 YNNKTDIWSLGCVLVEYELCTLKHPEGNLQOLVLIQAHFAPISPGFSRELHSLISOLF 240
Db 181 YNNKTDIWSLGCVLVEYELCTLKHPEGNLQOLVLIQAHFAPISPGFSRELHSLISOLF 240
QY 241 QVSPDRPSINSILKRPFLNLIPIKYLTPVIOEFSHMLICRAGPASPASHAGKVVQCK 300
Db 241 QVSPDRPSINSILKRPFLNLIPIKYLTPVIOEFSHMLICRAGPASPASHAGKVVQCK 300
QY 301 IQKVRFRGKCPPRSIRSVPIKRNAILHRNEWPPAGAKARSIMIERPKIAAVCGHYD 360
Db 301 IQKVRFRGKCPPRSIRSVPIKRNAILHRNEWPPAGAKARSIMIERPKIAAVCGHYD 360
QY 361 YYAQLDMLRRRAHPSYHP1PQENTGVEDYQETRHGSPSQWPAEYLRKFAQQYK 420
Db 361 YYAQLDMLRRRAHPSYHP1PQENTGVEDYQETRHGSPSQWPAEYLRKFAQQYK 420
QY 421 VEKQLGLRPSAEPNYNQRELRSNGBEPRFQELPFRKNEMKEQYWKQLEERQQVYND 480

Db 421 VEKQLGLRPSAEPNYNQRELRSNGBEPRFQELPFRKNEMKEQYWKQLEERQQVYND 480
QY 481 MKEIRKMGREPEDIKDLKQMLQNTKESKNPEQYKAKKGKVFPEINLDCISDENILQ 540
Db 481 MKEIRKMGREPEDIKDLKQMLQNTKESKNPEQYKAKKGKVFPEINLDCISDENILQ 540
QY 541 EEEAMDIPNEILTTFEDGKMKFKEVCVKEHGDYTDKAFKHLCPAAFTETLWLSFLFLEY 600
Db 541 EEEAMDIPNEILTTFEDGKMKFKEVCVKEHGDYTDKAFKHLCPAAFTETLWLSFLFLEY 600
QY 601 SLPHFLLEKSPFSRHLLIEDLLCANDCSLKDWSEKEMELRT 640
Db 601 SLPHFLLEKSPFSRHLLIEDLLCANDCSLKDWSEKEMELRT 640
RESULT 3
US-09-873-404-4
; Sequence 4, Application US/09873404
; Patent No. 6500656
; GENERAL INFORMATION:
; APPLICANT: WEBSTER, Marion et al
; TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC
; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES
; TITLE OF INVENTION: THEREOF
; FILE REFERENCE: CL001212-CIP
; CURRENT APPLICATION NUMBER: US/09/873,404
; CURRENT FILING DATE: 2001-06-05
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 616
; TYPE: PRT
; ORGANISM: Mouse
US-09-873-404-4
Query Match 31.2%; Score 1062.5; DB 4; Length 616;
Best Local Similarity 36.7%; Pred. No. 3.5e-79;
Matches 246; Conservative 108; Mismatches 148; Indels 169; Gaps 18;
QY 1 MDKYDV1KAIQGGAGFKAYLAKGSDSKHCVIKKEINFEKMPIQEKASKEVILEKMKH 60
Db 1 MEKYVRLQKIGEGFGKAVLVKSTEDGRHYVIKKEINSRMSDKERQSRREVAVLANMKH 60
QY 61 PNIVAFNPSFQENGRLEFVMEYCDGGDLMKRINRQGVLFSEDOILGWFWQISLGLKH 120
Db 61 PNIVQKESFEENGSLYVMDYCEGDLFKRINAKGALFOEDQILDWFWQICLALKHVH 120
QY 121 DRKILHRDIKAQNTFLSKNGVAKLGFGIARVLNNSMELARTCIGTPPYLSPEICQNK 180
Db 121 DRKILHRDIKAQNTFLSKNGVAKLGFGIARVLNNSMELARTCIGTPPYLSPEICQNK 180
QY 181 YNNKTDIWSLGCVLVEYELCTLKHPEGNLQOLVLIQAHFAPISPGFSRELHSLISOLF 240
Db 180 YNNKTDIWSLGCVLVEYELCTLKHPEGNLQOLVLIQAHFAPISPGFSRELHSLISOLF 240
QY 241 QVSPDRPSINSILKRPFLNLIPIKYLTPVIOEFSHMLICRAGPASPASHAGKVVQCK 285
Db 240 KENPRDRPSVNSILKRGFIKRIKFLSPQIABEFCLKTLSEKFGPQLPGKRPASQGV 299
QY 286 -----APASHAGKVVQCKIOKVRFRGKCPPRSIRSVPIKRNAILHRNEWPP 333
Db 300 SSVFPAQKITPAKYGVPLYIKYKGLLEKPPGKQAHQIPVK-----MNSGERK 356
QY 334 PAG--AQKARSIMIER----- 348
Db 357 KMSAEAKKRLEFIEKKQKQDQIRFLKAEQMKRQEKRLERINRAEQGRNVLRAGG 416
QY 349 -----PKIAAVCGHYDYVAQLDMLRR-RAH-----KPSYHP--IPOE 383
Db 417 GSEVKASFFGIGGAVSPSPCPRGQYEHYHAI PQMQLRAEDNEARKWGIYGNWLP 476
QY 384 NTG---VEDYQETRHGSPSQWPAEYLRKFAQQYK-----KQLGRPSS 431

Db 477 QKGLAVERANQ-----VBEFLQKREAMQKARABGHVYVYLARLQRIQLQ--- 522
Qy 432 AEPNYNORQLRNGSEEPFQBLPFRKNMKEQYWKQLEIRQOYLNDMKEIRKMGRE 491
Db 523 ---NFEROQIKAK-----LRGENKEADGTK-----GOE 548
Qy 492 PEDIEKDLQMRLONTKESKNP-----BQYKAKKGKVKFEINLDCISDENILOEBEAMD 546
Db 549 ATE-ETDMRLKKVESLKAQTNARAVALKEQLERKKEAYE--REKKWSEHLVARVKSSD 605
Qy 547 IPNETLTFEDG 557
Db 606 VPLPLELLEGTG 616

RESULT 4

US-10-243-735-4
; Sequence 4, Application US/10243735
; Patent No. 6706510
; GENERAL INFORMATION:
; APPLICANT: WEBSTER, Marion et al
; TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC
; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES
; FILE REFERENCE: CL001212DIV
; CURRENT APPLICATION NUMBER: US/10/243,735
; CURRENT FILING DATE: 2002-09-16
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 616
; TYPE: PRT
; ORGANISM: Mouse
US-10-243-735-4

Query Match 31.2%; Score 1062.5; DB 4; Length 616;
Best Local Similarity 36.7%; Pred. No. 3.5e-79;
Matches 246; Conservative 108; Mismatches 148; Indels 169; Gaps 18;

Qy 1 MDKYDVIKAGGAFKAYLAKGSKDKEHCVIKEINFEKMPIOEKEASKKEVILLKMKH 60
Db 1 MEKYVRLQKIGSGFKAVLKVSTEDGHRVYIKEINISMSDKERQESRRREVAVLANMKH 60
Qy 61 PNIVAFNFSQENGRLFIWMEYCDGDLMKRINRQGVLFSDQILGWFOISLGLKHH 120
Db 61 PNIVQKESFEENGSLIWMYCEGDLFKRNAQKALFQSDQILDWFWQICLALKEVH 120
Qy 121 DRKILHRDIKAONIPLSKNGMVAKLGDGFIARVANNMELARTCTGTPYLSPEICNKP 180
Db 121 DRKILHRDIKSONIFLTKDGTV-QLGDFGIARVANNMELARTCTGTPYLSPEICNKP 179
Qy 181 YNNKTDIWSLGCVLVELCTLKHFPFGNNLQVLKICQAHFAPISPGFSRELHSLISOLF 240
Db 180 YNNKSDIWMALGCVLYELCTLKHAFAGNMKNVLVXIISGSPFPVSPHYSDLSLSOLF 239
Qy 241 QVSPDRPDSINSILKRPFLNLIPIKYLTPVTOEBSHMLICRAG----- 285
Db 240 KENPRDRPDSVNSILEKGFIAKRIEFLSPOLIAEFCFLKTSKFGPQPLGKRPAAGGV 299
Qy 286 -----APASHRAGKVQCKIKQVPRGKCPPRSIS--VPIKRNAILHNEWRP 333
Db 300 SEFVPAQKITKPAKYGVPLTKYKYGDKLLEKPPPHKQAHQIPVK--MNSGEERK 356
Qy 334 PAG--AQKARSIMIER----- 348
Db 357 KMSEEAARKRLLEFIEKEKKQDQIRFLKAEQKQKQRLERINRAREQGWNVLRAGG 416
Qy 349 -----PKTAAVGHVYVYQALDMLRR-RAH-----KPSVHP--IQE 383
Db 417 SEVKASPTGIGVAGVSPSCSPRGQVHYHAIFDQMQRURADNEARWKGGIYGRWLPER 476
Qy 384 NTG---VEDYGOETRHGSPSQWPAEYLQKPEAQYKLVKVE-----KQLGRPSS 431

Db 477 QKGLAVERANQ-----VBEFLQKREAMQKARABGHVYVYLARLQRIQLQ--- 522
Qy 432 AEPNYNORQLRNGSEEPFQBLPFRKNMKEQYWKQLEIRQOYLNDMKEIRKMGRE 491
Db 523 ---NFEROQIKAK-----LRGENKEADGTK-----GOE 548
Qy 492 PEDIEKDLQMRLONTKESKNP-----BQYKAKKGKVKFEINLDCISDENILOEBEAMD 546
Db 549 ATE-ETDMRLKKVESLKAQTNARAVALKEQLERKKEAYE--REKKWSEHLVARVKSSD 605
Qy 547 IPNETLTFEDG 557
Db 606 VPLPLELLEGTG 616
RESULT 5
US-09-173-581-6
; Sequence 6, Application US/09173581A
; Patent No. 6013455
; GENERAL INFORMATION:
; APPLICANT: Bandman, Olga
; APPLICANT: Tang, Y. Tom
; APPLICANT: Hillman, Jennifer L.
; APPLICANT: Yue, Henry
; APPLICANT: Guegler, Karl J.
; APPLICANT: Corley, Neil C.
; APPLICANT: Gorgone, Gina
; APPLICANT: Azimzai, Yalda
; APPLICANT: Lu, Aina
; TITLE OF INVENTION: Protein Kinase Homologs
; FILE REFERENCE: PP-0614 US
; CURRENT APPLICATION NUMBER: US/09/173,581A
; CURRENT FILING DATE: 1998-10-15
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: PERL Program
; SEQ ID NO 6
; LENGTH: 345
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE: -
; OTHER INFORMATION: 1567782
US-09-173-581-6

Query Match 24.0%; Score 815.5; DB 3; Length 345;
Best Local Similarity 45.1%; Pred. No. 3.7e-59;
Matches 173; Conservative 56; Mismatches 80; Indels 75; Gaps 7;

Qy 1 MDKYDVIKAGGAFKAYLAKGSKDKEHCVIKEINFEKMPIOEKEASKKEVILLKMKH 60
Db 1 MEKYVRLQKIGSGFKAILVKSTEDGHRVYIKEINISMSDKERQESRRREVAVLANMKH 60
Qy 61 PNIVAFNFSQENGRLFIWMEYCDGDLMKRINRQGVLFSDQILGWFOISLGLKHH 120
Db 61 PNIVQYRESF-----EGILDWFWQICLALKEVH 88
Qy 121 DRKILHRDIKAONIPLSKNGMVAKLGDGFIARVANNMELARTCTGTPYLSPEICNKP 180
Db 89 DRKILHRDIKSONIFLTKDGTV-QLGDFGIARVANNMELARTCTGTPYLSPEICNKP 147
Qy 181 YNNKTDIWSLGCVLVELCTLKHFPFGNNLQVLKICQAHFAPISPGFSRELHSLISOLF 240
Db 148 YNNKSDIWMALGCVLYELCTLKHAFAGNMKNVLVXIISGSPFPVSPHYSDLSLSOLF 207
Qy 241 QVSPDRPDSINSILKRPFLNLIPIKYLTPVTOEBSHMLICRAG---PASHAG--- 293
Db 208 KENPRDRPDSVNSILEKGFIAKRIEFLSPOLIAEFCFLKTSKFGSQPIPAKRPASQNS 267
Qy 294 -KVQCKIKQVPRGKCPPRSISVPI-----KRNAILHNEWRP----- 334
Db 268 ISVMPAQKITK-----PAKYGVPLTKYKYGDKLLEKPPPHKQAHQIPKEKVT 319
Qy 335 -----AGAOKARSIMIERPK 350

Db 320 GEERRKISEEAAKRRLEFIEKDK 343

RESULT 6

US-09-420-915-6

; Sequence 6, Application US/09420915

; Patent No. 6264947

; GENERAL INFORMATION:

; APPLICANT: Bandman, Olga

; APPLICANT: Tang, Y. Tom

; APPLICANT: Hillman, Jennifer L.

; APPLICANT: Yue, Henry

; APPLICANT: Guegler, Karl J.

; APPLICANT: Corley, Neil C.

; APPLICANT: Gorgone, Gina

; APPLICANT: Azimzai, Valda

; APPLICANT: Lu, Aina

; TITLE OF INVENTION: Protein Kinase Homologs

; FILE REFERENCE: PF-0614 US

; CURRENT APPLICATION NUMBER: US/09/420,915

; CURRENT FILING DATE: 1999-10-20

; EARLIER APPLICATION NUMBER: US 09/173,581

; EARLIER FILING DATE: 1998-10-15

; NUMBER OF SEQ ID NOS: 18

; SOFTWARE: PERL Program

; SEQ ID NO 6

; LENGTH: 345

; TYPE: PRT

; ORGANISM: Homo sapiens

; FEATURE:

; OTHER INFORMATION: 1567782

US-09-420-915-6

Query Match

Best Local Similarity 24.0%; Score 815.5; DB 3; Length 345;

Matches 173; Conservative 56; Mismatches 80; Indels 75; Gaps 7;

Qy 1 MDKVDVTKAQGAFKAYLAGKSDSKVKEINFEKMPQIOEKASKKEVILLKMKH 60

Db 1 MEKVRLQKIGEGFGKAILVKSTEDGRQYVKEINISRMSSKERESREAVLANNKH 60

Qy 61 PNIVAFNFSQENGRLEFIVMEYCDGDLMKRINRQGVLPSEDIQWVFOISLGLKHIH 120

Db 61 PNIVQYRESP-----EGILDWVFOICLAKVH 88

Qy 121 DRKILHRDIAQNIIFLKNQVAKLGDGFIARVLNNSMELARTCIGTPYYLSPEICNKP 180

Db 89 DRKILHRDIAQNIIFLTKDGTV-QLGDFGIARVLNSTVELARTCIGTPYYLSPEICENKP 147

Qy 181 YNNKTDLSLGCVLVELCTLKHPEGNLQOLVLIKQAHFAPISPCFSRELHSLISQLF 240

Db 148 YNNKSDIWLGCVLVELCTLKHPEAGSMKNLVLIKISGSPFVSLHYSLRSLVSLF 207

Qy 241 QVSPDRPSINSILKRPLENIPKYLTPVIOEFESHMLICRAGA---PASRHAG---- 293

Db 208 KRNPRDRPSVNSILEKGFIAKRIEKLPSQLIAEFCFKTFSKFGSQPIPAKRPASQNS 267

Qy 294 -KVQKCKIQVFRGKCPPRSISVPI-----KRNAILHRNWRP----- 334

Db 268 ISVPAQKIKT-----PAKYGIPLAYKYGDKLHKPLQKHQAHTPEKRVNT 319

Qy 335 -----AGAAQKARSIMIERPK 350

Db 320 GEERRKISEEAAKRRLEFIEKDK 343

RESULT 7

US-09-538-092-1182

; Sequence 1182, Application US/09538092

; Patent No. 6753314

; GENERAL INFORMATION:

; APPLICANT: Giot, Loic

; APPLICANT: Mansfield, Traci A.

; TITLE OF INVENTION: Protein-Protein Complexes and Method of Using Same

; FILE REFERENCE: 15966-542

; CURRENT APPLICATION NUMBER: US/09/538,092

; CURRENT FILING DATE: 2000-03-29

; PRIOR APPLICATION NUMBER: 60/127,352

; PRIOR FILING DATE: 1999-04-01

; PRIOR APPLICATION NUMBER: 60/178,965

; PRIOR FILING DATE: 2000-02-01

; NUMBER OF SEQ ID NOS: 1387

; SOFTWARE: CuraPatSeqFormatter Version 0.9

; SEQ ID NO 1182

; LENGTH: 459

; TYPE: PRT

; ORGANISM: Homo sapiens

; FEATURE:

; NAME/KEY: misc_feature

; LOCATION: (0)...(0)

; OTHER INFORMATION: Polypeptide Accession Number P51956

US-09-538-092-1182

Query Match

Best Local Similarity 21.5%; Score 731; DB 4; Length 459;

Matches 175; Conservative 79; Mismatches 151; Indels 112; Gaps 13;

Qy 49 KKEVILLKMKHPNIVAFNFSQENGRLEFIVMEYCDGDLMKRINRQGVLPSEDIQWV 108

Db 1 RKEAVLSARKMKHPNIVAFNFSQENGRLEFIVMEYCDGDLMKRINRQGVLPSEDIQWV 60

Qy 109 FVOISLGLKHIHDKILHRDIAQNIIFLKNQVAKLGDGFIARVLNNSMELARTCIGTP 168

Db 61 FTQCLGVNHIHKKVLRHDIKSNIFLTQNGKV-KLGDGFSARLLSNPMAFACTYVGP 119

Qy 169 YVLSPEICQNKPYNNKTDIWSLGCVLVELCTLKHPEGNLQOLVLIKQAHFAPISGPF 228

Db 120 YVVPPEIWEINLPPYNNKSDIWSLGCVLVELCTLKHPEGNLQOLVLIKQAHFAPISGPF 179

Qy 229 SRELHSLISQLFQVSPDRPSINSILKRPLENIPKYLTPVIOEFESHMLICRAGAPA 288

Db 180 SYELOFLVQKQKPNRSHRPSATLLSRGIVARLVQKCLPPEIIVE----- 225

Qy 289 SRHAGKVQKCKIQVFRGKCPPRSISVPIKRNAILHRNWRPAPAGAKARSIMIER 348

Db 236 --YGEVLEEIKNSKHNTPRKKTNPRIIRIALGNEASTVQEEQCRKGS----- 272

Qy 349 PKIAAVCHYDYVYAAQLDMLRRRAKPSYHPIPOENTGVEDYGOETRHGSPS----QWP 404

Db 273 -----HTDLESINENLVESALRRVNEEKG--NKSVHLRRKASPNLHRRQW- 316

Qy 405 ABYLQKFEAQYKQKLVKVEKQLGRPSSAEPNPNQRBLSRN--GEEPR-FOELPFRKNEM 461

Db 317 -----EKNV---PNTALTALENASILTSLTAEDDRGGSGVSKNTT 356

Qy 462 KEQEVWKLSEIRQYLNDMKE-----IRKMGRE----- 491

Db 357 RKQ--W--LKETPTDLNLKADLSLAFQTYTYRPFSEGLFKPLSEATEASDVGG 412

Qy 492 -----PEDTEKDLQWRLQNTKESKNQKQYKAKK 521

Db 413 HDSVILDFERLEPGLDEEDTDFEEDDNDPQWSELKK 449

RESULT 8

US-09-538-092-1183

; Sequence 1183, Application US/09538092

; Patent No. 6753314

; GENERAL INFORMATION:

; APPLICANT: Giot, Loic

; APPLICANT: Mansfield, Traci A.

; TITLE OF INVENTION: Protein-Protein Complexes and Method of Using Same

; FILE REFERENCE: 15966-542

; CURRENT APPLICATION NUMBER: US/09/538,092

; CURRENT FILING DATE: 2000-03-29

; PRIOR APPLICATION NUMBER: 60/127,352

```

; PRIOR FILING DATE: 1999-04-01
; PRIOR APPLICATION NUMBER: 60/178,965
; PRIOR FILING DATE: 2000-02-01
; NUMBER OF SEQ ID NOS: 1387
; SOFTWARE: CurapatSeqFormatter Version 0.9
; SEQ ID NO 1183
; LENGTH: 841
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (0)..(0)
; OTHER INFORMATION: Polypeptide Accession Number P51957
US-09-538-092-1183

Query Match          19.2%; Score 655; DB 4; Length 841;
Best local Similarity 24.0%; Pred. No. 2.3e-45;
Matches 195; Conservative 136; Mismatches 251; Indels 230; Gaps 22;

Qy 4 YDVIKAGCGAFGKAYLAKGSKDSKHCIVKEINFEKMPIOEKSAKKEVILLSEKMPHNI 63
Db 6 YCIVLRVVGKSYGEVILVKRRDCKQVVIKLNLRNASSRRERRAAEQALLSQDKHPNI 65

Qy 64 VAFNSFQ-ENGRLFIVMEYCDGDLMKRINRQGVLFSEDDQLGMFVQISLGLKHIHDR 122
Db 66 VTVKESWEGDGLLIYVMGFCGGDLRYLKEQGQLLPENQVVFVQIAMAALQYLHEK 125

Qy 123 KILHRDIKAGNIPLSKNGMTAKLDFGIAEVLNNSMELASTCTGTPYLSPEICQNPYN 182
Db 126 HILHRDUKTONVFLTRNII-KYVDLGIARVLENHCHDMASTLTGTPYVMSPELFSNKPYN 184

Qy 183 NKTDIWSLGGVLYELCTLPHPFGNNILQQLVLTICQAHFAPISFGFSREHLHSLISOLFQV 242
Db 185 YKSDVMALGCCVYEMATLKHAFNAKDNWSLVYRIIEGKLPAAMPDYSPELAEIIRTMLSK 244

Qy 243 SPDRPNSILKRPFLLENLIPKYL- 268
Db 245 RPBERPSVRSLRQPYIKRQISFFLEATKIKTSKNINIKNGDSQSKPPFATVVSGEAESNHE 304

Qy 269 ---PEVIOEFSHMLICRAG- 298
Db 305 VIHPQLSSSGSQTYINGEKCLSQEKPRASGULLKSPASLKAHTCKDLSNTTELATISS 364

Qy 299 -----CKIQKV-----RFRGKCQPPSR 315
Db 365 VNIDILPAKGRDSVSGFVQENOPRYLDASNELGIGCISQVEEEMLQDNTKSSAOPENL 424

Qy 316 ISVPIKENAIL--HRNE---WPPAGAQAARSIKMIERP- 350
Db 425 I--PMWSSDITVTEKNEFVFKPFLPIKEQPKQOSLALSPKLSCSGTILAHSNLRLLGSS 482

Qy 351 -----IAACVGHYDYTYAQLDMLRRAHKPSYHP-IPQENTGVDEYDQETRHGP 398
Db 483 DSPASARVAGITGVCHHAQDVAGECIIIEKQGR---IHPDLQPHNSGSE-----P 530

Qy 399 SPQWPAEYIQRKEAQQYKLYEKQIGL---RPSSAEPNY-----NOR 439
Db 531 SLSRQRQKREOTEHRGEKQRVRDLFAFQESFPFLPSHPHIVGVVDVTSTOKEAENOR 590

Qy 440 QELRSNGEERFQELPRPKNMEKEQYVWQLEIRIQOYLNDMKEIRKIM-----GREPED 494
Db 591 RVTGVSSSRSSESMSSSKDRPLSARERRRLKQSQEEMSSSGSVKASISVAGPKPQE 650

Qy 495 IEKDLKQML- 532
Db 651 EDQPLFARRLSDCSVTQERKQIHCLSEDELSSTSTSDKSDGDIYGGKQGTNEINALVQ 710

Qy 533 ISDENIL--DEEEMADIP--NETLTFEDGKMFKEYEYCKEHGVDYDKAFELKHCPEAAFT 588
Db 711 LMTQTLKLDSEKESCEDVPVANFVSEFKLHKRYDRTLIL--HGKVAEBA-EEIHFKE---- 763

Qy 589 ELTWLSFLFELSPLPHFLFEKSPFSRHIEDL 620

```

Db 764 -----LPSAINPGSEKIRRLVEVL 782

RESULT 9
US-09-992-481-2
; Sequence 2, Application US/09992481
; Patent No. 6593125
; GENERAL INFORMATION:
; APPLICANT: Friddle, Carl Johan
; APPLICANT: Hilbun, Erin
; APPLICANT: Mathur, Brian
; APPLICANT: Turner, C. Alexander Jr.
; TITLE OF INVENTION: No. 6593125el Human Kinases and Polynucleotides
; FILE REFERENCE: LEX-0266-USA
; CURRENT APPLICATION NUMBER: US/09/992,481
; CURRENT FILING DATE: 2001-11-19
; PRIOR APPLICATION NUMBER: US 60/252,011
; PRIOR FILING DATE: 2000-11-20
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 692
; TYPE: PRT
; ORGANISM: homo sapiens
US-09-992-481-2

Query Match 18.7%; Score 638; DB 4; Length 692;
Best Local Similarity 39.3%; Pred. No. 4.3e-44;
Matches 148; Conservative 62; Mismatches 128; Indels 40; Gaps 8;

QY	1	MDKYDVIKAI	GQAFGKAYL	AKGSDSKHCV	IKKEINPEK	MPIOEK	EASKEVILL	EKWKH	60
DB	1	MEKYEIRV	VVGGA	FVHLCL	KAQKLV	IKQIP	VEQMTKE	ERQAQNEC	QVLLKLNH 60
QY	61	PNIVAFN	SPQENGR	LFI	VMEYCD	GGDLMKR	INRQGV	LSE	QOILGW
DB	61	PNVIEY	ENFLED	KALMI	AMEYAP	GGTFL	AEFIQ	KRCNSL	EEETILHFFVQ
QY	121	DRKILH	RD	KAQNI	FLSKXG	MAKLD	GF	GIARVL	NNSMELAR
DB	121	THULLH	RD	LKTQ	NILL	DKRWV	VVIG	DG	ISKIL-SKSKAY
QY	181	YNNKTD	IS	LGCV	LYEL	CTL	KHPF	EGNN	LOQLVK
DB	180	YNQKSD	I	WALG	CVLY	ELAS	LKRA	FEAAN	PALVLK
QY	241	QVSPR	DP	BS	INS	TL	KPFL	ENL	PKVLT
DB	240	SLS	PAQ	RP	PLSH	IAQ	PL	CI	RALN
QY	299	--CKIO	KV	FRG	--KCP	RS	RIS	VP	T
DB	286	TGSR	TT	SV	RC	GP	IP	GR	VP
QY	341	---	RS	KMI	--E	R	P	KIAA	353
DB	346	GV	TR	S	GR	L	I	WE	A

RESULT 10
US-10-196-927-2
; Sequence 2, Application US/10196927
; Patent No. 6797510
; GENERAL INFORMATION:
; APPLICANT: Walke, D. Wade
; APPLICANT: Scoville, John
; APPLICANT: Friddle, Carl Johan
; TITLE OF INVENTION: No. 6797510el Human Kinases and Polynucleotides Encoding
; FILE REFERENCE: LEX-0348-USA
; CURRENT APPLICATION NUMBER: US/10/196,927
; CURRENT FILING DATE: 2002-05-20
; PRIOR APPLICATION NUMBER: US 60/293,248

```
; PRIOR FILING DATE: 2001-05-24
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 645
; TYPE: PRT
; ORGANISM: homo sapiens
US-10-196-927-2

Query Match      17.3%; Score 590; DB 4; Length 645;
Best Local Similarity 28.4%; Pred. No. 3.6e-40;
Matches 161; Conservative 106; Mismatches 197; Indels 102; Gaps 17;

QY 3 KYDVKAIGQAGKAYLA---KKSQSKHCIVKEINFEKMPIOEKASKEVILLKMK 59
Db 28 RYVLOQKLGSGFTVYLVSDKAKRGEELKVLKEISVGLNPNETVQANLEAQLSKLD 87
QY 60 HPNTVAFNSFOENGRLFIWVEYCDGDLMKRIN--RQGVLFSEDOILGWFOISLGLK 117
Db 88 HPAIVKHFASVEQDNFCIITEYCEGRDLDDKIQEYKQAGKIPENQIIEFWIQLLQVD 147
QY 118 HIHDKILHRDIKAQNIIFLSKNGMVAKLGDGFIARVLNNSMELARTCIGTPYVLSPEICQ 177
Db 148 YMHERRILHRDLKSKNVLKNN--LLKIGDFGVSRLLMGSCDLATTTLTGTPHYMSPEALK 205
QY 178 NKPYNNKTDIWSLGCVLVELCTLKHPPGNNLQOLVKICQAHFAPISPGFSRELHSLIS 237
Db 206 HOGYDTRSDIWSLACILYEMCCMHAFAGSNFLSIVLKIVEGDTSPUPERYPKELNAIME 265
QY 238 QLFQVSPDRPSINSILKRPF---LENLIPKY--LTPE---VIOEFSHMLICRAGAP 287
Db 266 SMLNKNPSLRSATEILKIPYLDQQLNLMCRYSEMTELEDKNLDCQKEAAHII----- 318
QY 288 ASRHAGKVQCKTKQVFRG---KCPPRSISVPIKRNAILHNRWPPAGAKARSIK 344
Db 319 ---NAMQKRIHIQTLRALSEVQKMTFRERMLR-----KLQAADEKARKLK 361
QY 345 MIERPKIAAVCHDYVYVYQALDMLRRRAHKPSYHPIPOENTGVEDYQOETRHGSPSQWP 404
Db 362 KIVEEK-----YEENSKRMQELSRNFQOLSVDVLHEKHLKG----- 399
QY 405 AYLQRFKEAQYKLVKVEKQLGLRPSAEPNYNORQELRNGEBEPFOELPFRKNEKEQ 464
Db 400 ---MEEKEQPEGLSCSPQ-----DEBERWQGRE---ESDEPTLENLIP----- 439
QY 465 EYWKOLBEIRQOYLNDMKEIRKMGREPEDIKQLQMLQNTKESKNPQYKAKKGVK 524
Db 440 ---ESQPIPSMDLHELESI-----VEDATSDLGYHEI---PEDPLVABEYVADAFDS 485
QY 525 FEINLKDICSDENILQEEAWDIPNE 550
Db 486 YCVESDE--BEEETALERPEKEIRNE 509

RESULT 11
US-10-196-927-4
; Sequence 4, Application US/10196927
; Patent No. 6797510
; GENERAL INFORMATION:
; APPLICANT: Walke, D. Wade
; APPLICANT: Scoville, John
; APPLICANT: Friddle, Carl Johan
; TITLE OF INVENTION: No. 6797510el Human Kinases and Polynucleotides Encoding the Same
; FILE REFERENCE: LEX-0348-USA
; CURRENT APPLICATION NUMBER: US/10/196,927
; CURRENT FILING DATE: 2002-05-20
; PRIOR APPLICATION NUMBER: US 60/293,248
; PRIOR FILING DATE: 2001-05-24
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 482
; TYPE: PRT
```

```
; ORGANISM: homo sapiens
US-10-196-927-4

Query Match      17.2%; Score 584; DB 4; Length 482;
Best Local Similarity 29.2%; Pred. No. 7.5e-40;
Matches 149; Conservative 94; Mismatches 164; Indels 104; Gaps 14;

QY 3 KYDVKAIGQAGKAYLA---KKSQSKHCIVKEINFEKMPIOEKASKEVILLKMK 59
Db 28 RYVLOQKLGSGFTVYLVSDKAKRGEELKVLKEISVGLNPNETVQANLEAQLSKLD 87
QY 60 HPNTVAFNSFOENGRLFIWVEYCDGDLMKRIN--RQGVLFSEDOILGWFOISLGLK 117
Db 88 HPAIVKHFASVEQDNFCIITEYCEGRDLDDKIQEYKQAGKIPENQIIEFWIQLLQVD 147
QY 118 HIHDKILHRDIKAQNIIFLSKNGMVAKLGDGFIARVLNNSMELARTCIGTPYVLSPEICQ 177
Db 148 YMHERRILHRDLKSKNVLKNN--LLKIGDFGVSRLLMGSCDLATTTLTGTPHYMSPEALK 205
QY 178 NKPYNNKTDIWSLGCVLVELCTLKHPPGNNLQOLVKICQAHFAPISPGFSRELHSLIS 237
Db 206 HOGYDTRSDIWSLACILYEMCCMHAFAGSNFLSIVLKIVEGDTSPUPERYPKELNAIME 265
QY 238 QLFQVSPDRPSINSILKRPF---LENLIPKY--LTPE---VIOEFSHMLICRAGAP 287
Db 266 SMLNKNPSLRSATEILKIPYLDQQLNLMCRYSEMTELEDKNLDCQKEAAHII----- 318
QY 288 ASRHAGKVQCKTKQVFRG---KCPPRSISVPIKRNAILHNRWPPAGAKARSIK 344
Db 319 ---NAMQKRIHIQTLRALSEVQKMTFRERMLR-----KLQAADEKARKLK 361
QY 345 MIERPKIAAVCHDYVYVYQALDMLRRRAHKPSYHPIPOENTGVEDYQOETRHGSPSQWP 404
Db 362 KIVEEK-----YEENSKRMQELSRNFQOLSVDVLHEKHLKG----- 399
QY 405 AYLQRFKEAQYKLVKVEKQLGLRPSAEPNYNORQELRNGEBEPFOELPFRKNEKEQ 458
Db 400 ---MEEKEQPEGLSCSPQ-----DEBERWQGRE---ESDEPTLENLIPESQPIPSMD 448
QY 459 -----NENKEQOEYWK 468
Db 449 LHELESIVEDATSDLGHGDCNLSLDEYWK 479

RESULT 12
US-08-870-529-2
; Sequence 2, Application US/08870529
; Patent No. 6080557
; GENERAL INFORMATION:
; APPLICANT: Sims, John E.
; APPLICANT: Virca, G. Duke
; APPLICANT: Bird, Timothy A.
; APPLICANT: Anderson, Dirk M.
; TITLE OF INVENTION: IL-1/TNF-(-ACTIVATED KINASE (ITAK),
; NUMBER OF INVENTION: AND METHODS OF MAKING AND USING THE SAME
; NUMBER OF SEQUENCES: 18
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SEED AND BERRY LLP
; STREET: 6300 Columbia Center, 701 Fifth Avenue
; CITY: Seattle
; STATE: Washington
; COUNTRY: USA
; ZIP: 98104
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/870,529
; FILING DATE: 06-JUN-1997
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
```

[illegible][illegible]

```
Qy      288 ASRHAGKVQVKCKTKVFRGKCPPRSISVPITKNAILHNEHWPPAGAOK----- 339
         | : : | : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      295 KSDSPSVLSLKLTQLQ-----ERERALKAREERLEQEOELCVRELA 341
         | : : | : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy      340 -----ARSTKMTERPKIAVCGHYDYYAQLDMLRRAHKPSVHPTPQENTGWED 389
```


Job time : 43 secs

```
Db 342 EDKLARAENLKNYSLLKRFSLASNPPELLNLPSSVKKVH---FSGESKENI---- 394
Qy 390 YGQETHRGPSQWPAEYLOKFEAQO-----YKLKVEKQLGLR 428
Db 395 MRSENSESQTSKCKDLKRLHAAQRAQALSDEIKNYOLKSRQILGMR 445

RESULT 15
US-09-538-092-1181
; Sequence 1181, Application US/09538092
; Patent No. 6753314
; GENERAL INFORMATION:
; APPLICANT: Giot, Loic
; APPLICANT: Mansfield, Traci A.
; TITLE OF INVENTION: Protein-Protein Complexes and Method of Using Same
; FILE REFERENCE: 15966-542
; CURRENT APPLICATION NUMBER: US/09/538,092
; CURRENT FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: 60/127,352
; PRIOR FILING DATE: 1999-04-01
; PRIOR APPLICATION NUMBER: 60/178,965
; PRIOR FILING DATE: 2000-02-01
; NUMBER OF SEQ ID NOS: 1387
; SOFTWARE: CuratPatSeqFormatter Version 0.9
; SEQ ID NO 1181
; LENGTH: 445
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (0)-(0)
; OTHER INFORMATION: Polypeptide Accession Number P51955
US-09-538-092-1181
```

```
Query Match 15.7%; Score 535.5; DB 4; Length 445;
Best Local Similarity 29.7%; Pred. No. 6.8e-36;
Matches 140; Conservative 89; Mismatches 167; Indels 75; Gaps 13;

Qy 2 DKYDVIKAIIGGAFKAYLAKGSDSKHCYVKEINFEKMPIOKEASKEVILLESKHP 61
Db 6 EDEVLYTIGTGYGRCQKIRKSDGKILYMKELDYGMTAEKQMLVSEVNLRELKHP 65
Qy 62 NIVAFNSF--QENGRFIVWEYCDGDLMKRI---NRQGVLFSDQILGWFVQISLGL 116
Db 66 NIVRYDEIIDRTNTTIIYWEYCEGDLASVITKTKERQYL-DEEFVLRVMTQLTAL 124
Qy 117 KHIHDRK-----ILHRDIKQNIPLSKNGVAKLGDGFIARVLNNSMELARTCIGIPYYL 171
Db 125 KECHRRSDGGHTVLHRLDKPANVFLDGQNV-KLGDGFLARILNHDTSFAKTFVGTPTYM 183
Qy 172 SPEICQKNYNNKTDIWSLGVLYELCTLKHPPPEGNNLOQLVKICOAHFAPISPGFSRE 231
Db 184 SPEQNMMSVNEKSDIWSLGLLYELCALMPPTAFSQKELACKIRGKFRIPYRISDE 243
Qy 232 LHSLSIQLFQVSPDRPSINSILKRPFLNIPKYLTPVIOEFSHMLICRA---GAP- 287
Db 244 LNEIITRLMLNDYHRPSVEILENPLIADL-----VADEQRNLERRGRQLGEPE 294
Qy 288 ASRHAGKVQCKIQKVRFRGKCPFRSISVPIKRNAILHNEWRPPAGAOK----- 339
Db 295 KSQDSSPVLSELKLEIQLQ-----ERERALKAREERLEQKEQLCVVERLA 341
Qy 340 -----ARSIKMIERPKIAVCGHYDYVAQLDMLRRRAHKPSYHPIQENTGVED 389
Db 342 EDKLARAENLKNYSLLKRFSLASNPPELLNLPSSVKKVH---FSGESKENI---- 394
Qy 390 YGQETHRGPSQWPAEYLOKFEAQO-----YKLKVEKQLGLR 428
Db 395 MRSENSESQTSKCKDLKRLHAAQRAQALSDEIKNYOLKSRQILGMR 445
```

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: November 12, 2004, 12:46:14 ; Search time 156 Seconds

(without alignments)
1471.710 Million cell updates/sec

Title: US-10-730-010-2

Perfect score: 3404

Sequence: 1 MDKYDVKAIGQAFGKAYL.....LCANDCSLKDWSEKEMELRT 640

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 2002273 seqs, 358729299 residues

Total number of hits satisfying chosen parameters: 2002273

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

- Database : A_Geneseq_23Sep04:*
- 1: Geneseq1980s:*
 - 2: Geneseq1990s:*
 - 3: Geneseq2000s:*
 - 4: Geneseq2001s:*
 - 5: Geneseq2002s:*
 - 6: Geneseq2003s:*
 - 7: Geneseq2003Bs:*
 - 8: Geneseq2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	3404	100.0	640	6	ABG72000 Human ser
2	3081.5	90.5	683	5	AAU77928 Aau77928 Amino aci
3	3081.5	90.5	889	8	ADJ96637 Adj96637 Human Nim
4	2945	86.5	654	5	AAU77929 Aau77929 Amino aci
5	2715.5	79.8	649	4	AAU03545 Aau03545 Human pro
6	2610.5	76.7	631	6	ABP71710 Abp71710 Human NEK
7	2107.5	61.9	403	6	ABU54635 Abu54635 Human NOV
8	2102.5	61.8	399	6	ABU54634 Abu54634 Human NOV
9	1840.5	48.2	614	7	ADF44528 Adf44528 Mouse kin
10	1135.5	33.4	336	7	ADF44528 Adf44528 Mouse kin
11	1108.5	32.6	544	6	ABP36073 Abp36073 Human pro
12	1086	31.9	1242	4	AAU39210 Aau39210 Human pol
13	1086	31.9	1242	6	AAU16440 Aau16440 Human ser
14	1078	31.7	774	6	AAO26613 Aao26613 Serine/th
15	1078	31.7	774	6	ABP71711 Abp71711 NEK-like
16	1078	31.7	774	6	AAO16441 Aao16441 Human ser
17	1078	31.7	774	6	ABP7691 Abp7691 Amino aci
18	1072	31.5	1286	8	ADJ96638 Adj96638 Human Nim
19	1060	31.1	1214	4	AAU07102 Aau07102 Human nov
20	1056	31.0	1214	4	AAU39211 Aau39211 Human pol
21	1056	31.0	1214	6	ABU08113 Abu08113 Human kin
22	928	27.3	632	5	ABP55160 Abp55160 Ser/Thr/T
23	820.5	24.1	506	6	ABU54631 Abu54631 Human NOV
24	820.5	24.1	506	6	ADH42547 Adh42547 Novel hum
25	815.5	24.0	345	3	AAU76753 Aau76753 Human pro

26	815.5	24.0	345	4	AAE06211 Aae06211 Human pro
27	815.5	24.0	345	5	ABB84438 Abb84438 Human pro
28	814	23.9	461	8	ADI40836 Adi40836 Human kin
29	812	23.9	425	6	ABU54633 Abu54633 Human NOV
30	812	23.9	425	8	ADH42551 Adh42551 Novel hum
31	812	23.9	489	6	ABU54632 Abu54632 Human NOV
32	812	23.9	489	8	ADH42549 Adh42549 Novel hum
33	810.5	23.8	506	4	AAU78344 Aau78344 Human pro
34	810.5	23.8	506	5	ABP60668 Abp60668 Human ser
35	810.5	23.8	506	8	ADJ96639 Adj96639 Human Nim
36	808.5	23.7	507	6	AAE24136 Aae24136 Human kin
37	808	23.7	507	6	ABP97688 Abp97688 Amino aci
38	806.5	23.7	527	4	AAU79328 Aau79328 Human pro
39	806.5	23.7	527	5	ABB97224 Abb97224 Novel hum
40	805.5	23.7	511	6	ABP71712 Abp71712 NEK-like
41	805.5	23.7	511	6	ABP97690 Abp97690 Amino aci
42	804.5	23.6	546	6	ABP96072 Abp96072 Human pro
43	802	23.6	510	8	ADM16429 Adm16429 Human kin
44	800	23.5	489	7	ADC93067 Adc93067 Human KPP
45	791.5	23.3	489	8	ADH42553 Adh42553 Novel hum

ALIGNMENTS

RESULT 1
ABG72000
ID ABG72000 standard; protein; 640 AA..

XX AC ABG72000;

XX XX 04-FEB-2003 (first entry)

XX DT Human serine/threonine kinase.

XX DE Human; enzyme; serine/threonine kinase; lung carcinoma; chromosome 13.

XX KW Homo sapiens.

XX OS WO200281727-A2.

XX PN PD 17-OCT-2002.

XX PF 02-APR-2002; 2002WO-US010156.

XX PR 03-APR-2001; 2001US-00824583.

XX PR 05-JUN-2001; 2001US-00873404.

XX XX (PEKE) PE CORP NY.

XX PI Webster M, Yan C, Di Francesco V, Beasley EM;

XX XX WPI; 2003-058562/05.

XX DR N-PSDB; ABS57149, ABS57150.

XX PT Novel human kinase protein expressed in lung carcinoma and placenta is useful to diagnose and treat diseases and disorders associated with expression or activity of the protein.

XX PS Claim 1; Fig 2A; 101pp; English.

XX CC The invention relates to an isolated human kinase peptide of the serine/threonine, an allelic variant or orthologue, and encoded by a nucleic acid that hybridises under stringent conditions with the cDNA and gene sequences appearing as ABS57149 and ABS57150, or a fragment comprising at least 10 contiguous amino acids. Also included are an antibody that selectively binds to the kinase, a gene chip comprising the nucleic acids, a transgenic non human animal comprising the nucleic acids, a nucleic acid vector comprising the nucleic acids, a host cell containing the vector and expressing the kinase, identifying a modulator/binding agent of the kinase (comprising contacting the peptide with an agent and determining if the agent has modulated function, expression or activity of the peptide or formed a complex with it), The

CC molecules of the invention are useful to diagnose and treat a disorder
CC characterised by aberrant expression of the protein (e.g. lung
CC carcinoma). Agents which modulate the function or activity of the protein
CC are useful to treat diseases or disorders mediated by human kinase
CC protein. The gene for encoding the novel kinase is located on chromosome
CC 13. The present sequence represents the novel human kinase
XX
SQ Sequence 640 AA;

Query Match 100.0%; Score 3404; DB 6; Length 640;
Best Local Similarity 100.0%; Pred. No. 1.1e-269;
Matches 640; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MDKVDVIAIGGAGFGKAYLAGKSDSKHCIVKEINFEKMPIQKEASKKEVILEKMKH 60
DB 1 MDKVDVIAIGGAGFGKAYLAGKSDSKHCIVKEINFEKMPIQKEASKKEVILEKMKH 60
QY 61 PNIVAFNFSFOENGRLFIVMEYCDGGDLMKRINRQGVLFSEDIQILGWFWQISLGLKHIH 120
DB 61 PNIVAFNFSFOENGRLFIVMEYCDGGDLMKRINRQGVLFSEDIQILGWFWQISLGLKHIH 120
QY 121 DRKILHRDIKAQNIFLSKNGMVAKLGDGFIARVLNNSMELARTICIGTPYVLSPEICQK 180
DB 121 DRKILHRDIKAQNIFLSKNGMVAKLGDGFIARVLNNSMELARTICIGTPYVLSPEICQK 180
QY 181 YNNKTDIWSLGCVLVELCTLKHPEGNLQOLVLKICQAHFAPISPGFSRHLHSLSOLF 240
DB 181 YNNKTDIWSLGCVLVELCTLKHPEGNLQOLVLKICQAHFAPISPGFSRHLHSLSOLF 240
QY 241 QVSPDRPSINSILKRPFLNLI PKYLTPEVIOEFSSHMLICRAGAPASRHAGKVQKCK 300
DB 241 QVSPDRPSINSILKRPFLNLI PKYLTPEVIOEFSSHMLICRAGAPASRHAGKVQKCK 300
QY 301 IOKVFRGCKPPRSRISVPIKRNAILHRNWRPAGAKARSIKMIERP KIAVCGHYD 360
DB 301 IOKVFRGCKPPRSRISVPIKRNAILHRNWRPAGAKARSIKMIERP KIAVCGHYD 360
QY 361 YYAQDMLRRRAHKPSYHIPQENTGVEDYQGETRHGSPSQWPAEYLQKFEAQYK 420
DB 361 YYAQDMLRRRAHKPSYHIPQENTGVEDYQGETRHGSPSQWPAEYLQKFEAQYK 420
QY 421 VEKQLGLRPSAEPNRYNORQLRNGEPPRQELPFRKNEMKEQYWKQLEERQOYLND 480
DB 421 VEKQLGLRPSAEPNRYNORQLRNGEPPRQELPFRKNEMKEQYWKQLEERQOYLND 480
QY 481 MKEIRKMGREPDIEKDLKQMLQNTKESKNPQKYKAKGVKFEINLDKCISDENILQ 540
DB 481 MKEIRKMGREPDIEKDLKQMLQNTKESKNPQKYKAKGVKFEINLDKCISDENILQ 540
QY 541 EEEAMDIPNETLTPEDGMKFEYCVKEHGDYTDKAFELKHCPEAAFTLWLSFLEY 600
DB 541 EEEAMDIPNETLTPEDGMKFEYCVKEHGDYTDKAFELKHCPEAAFTLWLSFLEY 600
QY 601 SLPHFLLEKSPFSLHIEDLLCANDCSLKDWSEKEMELRT 640
DB 601 SLPHFLLEKSPFSLHIEDLLCANDCSLKDWSEKEMELRT 640

RESULT 2
AAU77928
ID AAU77928 standard; protein; 683 AA.
XX AAU77928;
AC AAU77928;
XX
DT 18-JUN-2002 (first entry)
XX Amino acid sequence for novel human kinase protein #1.
DE
KW Novel human protein; NHP; serine-threonine kinase; brain;
KW calcium/calmodulin-dependent kinase; myosin light chain kinase;
XX biological disorder; spleen; placenta; chromosome 6; enzyme.
OS Homo sapiens.

XX WO200218555-A2.
PN 07-MAR-2002.
XX 28-AUG-2001; 2001WO-US026775.
XX 31-AUG-2000; 2000US-0229280P.
XX (LEXI-) LEXICON GENETICS INC.
PI Friddle CJ, Hilbun E, Nepomnichy B, Hu Y;
XX WPI: 2002-292200/33.
DR N-PSDB; ABK12424.
XX Novel polynucleotide encoding novel human protein sharing structural
PT similarity with animal kinases e.g. serine-threonine, calcium/calmodulin-
PT dependent, and myosin light chain kinases, useful as probes and primers.
XX Claim 1; Page 38-39; 46pp; English.
XX The present invention relates to the isolation of novel human proteins
CC (NHPs) and the polynucleotide sequences encoding them. The NHPs of the
CC invention are kinase proteins and share structural similarity to serine-
CC threonine, calcium/calmodulin-dependent, and myosin light chain kinases.
CC The sequences of the invention are useful for treating biological
CC disorders. The polynucleotide sequences encoding the kinase proteins can
CC be used as primers and probes. The sequences are also useful for
CC identifying mutations associated with a particular disease and also in a
CC prognostic or diagnostic assay. The present sequence represents human
CC protein kinase #1 which is expressed in a broad range of human tissues
CC such as brain, spleen, and placenta. The gene encoding protein kinase #1
CC is located on chromosome 6
SQ Sequence 683 AA;

Query Match 90.5%; Score 3081.5; DB 5; Length 683;
Best Local Similarity 94.1%; Pred. No. 3.1e-243;
Matches 585; Conservative 2; Mismatches 4; Indels 31; Gaps 1;
QY 1 MDKVDVIAIGGAGFGKAYLAGKSDSKHCIVKEINFEKMPIQKEASKKEVILEKMKH 60
DB 1 MDKVDVIAIGGAGFGKAYLAGKSDSKHCIVKEINFEKMPIQKEASKKEVILEKMKH 60
QY 61 PNIVAFNFSFOENGRLFIVMEYCDGGDLMKRINRQGVLFSEDIQILGWFWQISLGLKHIH 120
DB 61 PNIVAFNFSFOENGRLFIVMEYCDGGDLMKRINRQGVLFSEDIQILGWFWQISLGLKHIH 120
QY 121 DRKILHRDIKAQNIFLSKNGMVAKLGDGFIARVLNNSMELARTICIGTPYVLSPEICQK 180
DB 121 DRKILHRDIKAQNIFLSKNGMVAKLGDGFIARVLNNSMELARTICIGTPYVLSPEICQK 180
QY 181 YNNKTDIWSLGCVLVELCTLKHPEGNLQOLVLKICQAHFAPISPGFSRHLHSLSOLF 240
DB 181 YNNKTDIWSLGCVLVELCTLKHPEGNLQOLVLKICQAHFAPISPGFSRHLHSLSOLF 240
QY 241 QVSPDRPSINSILKRPFLNLI PKYLTPEVIOEFSSHMLICRAGAPASRHAGKVQKCK 300
DB 241 QVSPDRPSINSILKRPFLNLI PKYLTPEVIOEFSSHMLICRAGAPASRHAGKVQKCK 300
QY 301 IOKVFRGCKPPRSRISVPIKRNAILHRNWRPAGAKARSIKMIERP KIAVCGHYD 360
DB 301 IOKVFRGCKPPRSRISVPIKRNAILHRNWRPAGAKARSIKMIERP KIAVCGHYD 360
QY 361 YYAQDMLRRRAHKPSYHIPQENTGVEDYQGETRHGSPSQWPAEYLQKFEAQYK 420
DB 361 YYAQDMLRRRAHKPSYHIPQENTGVEDYQGETRHGSPSQWPAEYLQKFEAQYK 420
QY 421 VEKQLGLRPSAEPNRYNORQLRNGEPPRQELPFRKNEMKEQYWKQLEERQOYLND 480
DB 421 VEKQLGLRPSAEPNRYNORQLRNGEPPRQELPFRKNEMKEQYWKQLEERQOYLND 480

```
QY 481 MKEIRKMGREPE-----DIEKDLKQRLQNTKE 509
D 481 MKEIRKMGREPEENSKISHKTYLVKSKNLPVHQDASEGEAPVQDIEKDLKQRLQNTKE 540
QY 510 SKNPEQYKAKGVKFEINLDCISDENILOBEEAMDIPNETLTPEDCMKFKEVCVKEH 569
D 541 SKNPEQYKAKGVKFEINLDCISDENILOBEEAMDIPNETLTPEDCMKFKEVCVKEH 600
QY 570 GDYTDKAFKHLHCPEAAFTLT 591
D 601 GDYTDKAFKHLHCPEAGFSTQT 622

RESULT 3
ID ADJ96637 standard; protein; 889 AA.
AC ADJ96637;
XX
DT 06-MAY-2004 (first entry)
DE Human Nim-A related protein kinase NEK5 protein SeqID 94.
XX
KW kinase; human; tyrosine protein kinase; serine/threonine protein kinase;
KW PK; STK; gene therapy; cancer; immune-related disease;
KW cardiovascular disease; brain; neuronal associated disease; metabolic;
KW inflammatory disorder; cytostatic; neuroprotective; immunomodulator;
KW antiinflammatory; enzyme; Nim-A related protein kinase; NEK5.
XX
OS Homo sapiens.
OS 56.
XX
FN W02004006838-A2.
XX
PD 22-JAN-2004.
XX
PF 15-JUL-2003; 2003WO-US021730.
XX
PR 15-JUL-2002; 2002US-0395632P.
XX
PA (SUGS-) SUGEN INC.
XX
PI Whyte D, Manning G, Caenepeel S;
XX
DR WPI; 2004-122753/12.
XX
DR N-PSDB; ADJ96571.
XX
PT New nucleic acid molecule encoding a kinase polypeptide, useful for
PT preparing a composition for treating diseases or disorders, e.g., cancer,
PT or neurological, immunological or inflammatory disorders.
XX
PS Claim 1; SEQ ID NO 94; 366pp; English.
XX
CC This invention relates to a novel isolated, enriched or purified nucleic
CC acid molecule that encodes a kinase polypeptide. Specifically, it relates
CC to human tyrosine and serine/threonine protein kinases (PK's and STK's),
CC as well as protein kinase-like enzymes. The present invention describes
CC screening methods to identify agonists, antagonists and antibodies that
CC can be used to modulate the activity or function of the mammalian kinase
CC enzymes. As such, these compositions can be used for gene therapy
CC purposes to treat diseases or disorders including cancer, immune-related
CC diseases, cardiovascular disease, brain or neuronal associated disease,
CC metabolic and inflammatory disorders. Accordingly, they exhibit
CC cytostatic, neuroprotective, immunomodulator and antiinflammatory
CC activities. This polypeptide sequence is a human kinase protein sequence
CC of the invention.
XX
SQ Sequence 889 AA;
Query Match 90.5%; Score 3081.5; DB 8; Length 889;
Best Local Similarity 94.1%; Pred. No. 4.4e-243;
Matches 595; Conservative 2; Mismatches 4; Indels 31; Gaps 1;
```

```
QY 1 MDKYDVIIKAIQGAFGKAYLAKGKSDSKHCVKEINFEKMPIQEKASKEVILEKMKH 60
D 58 MDKYDVIIKAIQGAFGKAYLAKGKSDSKHCVKEINFEKMPIQEKASKEVILEKMKH 117
QY 61 PNIVAFNSFOENGRLEFVMEYCDGDLMKRIINRQGVLFSEQILGMFVQISLGLKHH 120
D 118 PNIVAFNSFOENGRLEFVMEYCDGDLMKRIINRQGVLFSEQILGMFVQISLGLKHH 177
QY 121 DRKILHRDIIKAQNIIFLSKNGWAKLGDGFIARVLNNSMELARTICGTPPYLSPEICQNK 180
D 178 DRKILHRDIIKAQNIIFLSKNGWAKLGDGFIARVLNNSMELARTICGTPPYLSPEICQNK 237
QY 181 YNNKTDIWSLGCVLVELCTLKHPEGNLQVLKICQAHFAPISPGFSELHSLISOLF 240
D 238 YNNKTDIWSLGCVLVELCTLKHPEGNLQVLKICQAHFAPISPGFSELHSLISOLF 297
QY 241 QVSPDRPSINSILKRPFLNLPKYLTPVIOBESHMLICRAGAPASHAGKVQCK 300
D 298 QVSPDRPSINSILKRPFLNLPKYLTPVIOBESHMLICRAGAPASHAGKVQCK 357
QY 301 IQKVRPRGKCPPRSIRISVPICKRNAILHRNEWPPAGAKARSIMIERPKIAAVCGHYD 360
D 358 IQKVRPRGKCPPRSIRISVPICKRNAILHRNEWPPAGAKARSIMIERPKIAAVCGHYD 417
QY 361 YYAOLDMLRRRAHKPSYHPIPOENTGVEDYQOETHGSPSPONPAEYLQKFEAQOYK 420
D 418 YYAOLDMLRRRAHKPSYHPIPOENTGVEDYQOETHGSPSPONPAEYLQKFEAQOYK 477
QY 421 VEKQLGLRPSAEPNPNQRLSNGEPRFQELPFRKNEMKEQYWKOLEEIRQQYLND 480
D 478 VEKQLGLRPSAEPNPNQRLSNGEPRFQELPFRKNEMKEQYWKOLEEIRQQYLND 537
QY 481 MKEIRKMGREPE-----DIEKDLKQRLQNTKE 509
D 538 MKEIRKMGREPEENSKISHKTYLVKSKNLPVHQDASEGEAPVQDIEKDLKQRLQNTKE 597
QY 510 SKNPEQYKAKGVKFEINLDCISDENILOBEEAMDIPNETLTPEDCMKFKEVCVKEH 569
D 598 SKNPEQYKAKGVKFEINLDCISDENILOBEEAMDIPNETLTPEDCMKFKEVCVKEH 657
QY 570 GDYTDKAFKHLHCPEAAFTLT 591
D 658 GDYTDKAFKHLHCPEAGFSTQT 679

RESULT 4
AAU77929
ID AAU77929 standard; protein; 654 AA.
AC AAU77929;
XX
DT 18-JUN-2002 (first entry)
DE Amino acid sequence for novel human kinase protein #2.
XX
KW Novel human protein; NHP; serine-threonine kinase; brain;
KW calcium/calmodulin-dependent kinase; myosin light chain kinase;
KW biological disorder; spleen; placenta chromosome 6; enzyme.
XX
OS Homo sapiens.
XX
PN W0200218555-A2.
XX
PD 07-MAR-2002.
XX
PF 28-AUG-2001; 2001WO-US026776.
XX
PR 31-AUG-2000; 2000US-0229280P.
XX
PA (LEXI-) LEXICON GENETICS INC.
XX
PI Friddle CJ, Hilbun E, Nepomnichy B, Hu Y;
XX
```

DR WPI; 2002-292200/33.

DR N-PSDB; ABK12425.

XX Novel polynucleotide encoding novel human protein sharing structural
PT similarity with animal kinases e.g. serine-threonine, calcium/calmodulin-
PT dependent, and myosin light chain kinases, useful as probes and primers.
XX
XX
PS Claim 3; Page 40-41; 46pp; English.

XX The present invention relates to the isolation of novel human proteins
CC (NHPs) and the polynucleotide sequences encoding them. The NHPs of the
CC invention are kinase proteins and share structural similarity to serine-
CC threonine, calcium/calmodulin-dependent, and myosin light chain kinases.
CC The sequences of the invention are useful for treating biological
CC disorders. The polynucleotide sequences encoding the kinase proteins can
CC be used as primers and probes. The sequences are also useful for
CC identifying mutations associated with a particular disease and also in a
CC prognostic or diagnostic assay. The present sequence represents human
CC protein kinase #2 which is expressed in a broad range of human tissues
CC such as brain, spleen, and placenta. The gene encoding protein kinase #2
CC is located on chromosome 6
XX
XX Sequence 654 AA;

Query Match

Best Local Similarity 86.5%; Score 2945; DB 5; Length 654;
Matches 564; Conservative 7; Mismatches 15; Indels 12; Gaps 4;

QY 1 MDKYDVIKAIQGGAGFKAYLAKGSDSKHCVIKENFEKMPIOKEASKKEVILEKWKH 60
DB 1 MDKYDVIKAIQGGAGFKAYLAKGSDSKHCVIKENFEKMPIOKEASKKEVILEKWKH 60
QY 61 PNIVAFNFSFQENGLFTWMEYCDGDLKMGINRQGVLFSEDIQILGFWQISLGLKHII 120
DB 61 PNIVAFNFSFQENGLFTWMEYCDGDLKMGINRQGVLFSEDIQILGFWQISLGLKHII 120
QY 121 DRKILHRDIKQNTIFLSKNGVAKLGFAGIARVLNNSMELARTCIGTPYLSPEICQNK 180
DB 121 DRKILHRDIKQNTIFLSKNGVAKLGFAGIARVLNNSMELARTCIGTPYLSPEICQNK 180
QY 181 YNNKTDIWSLGCVLVELCTLKHPEGNLQOLVLKI COAHFAPISPGFSRELHSLISOLF 240
DB 181 YNNKTDIWSLGCVLVELCTLKHPEGNLQOLVLKI COAHFAPISPGFSRELHSLISOLF 240
QY 241 QVSPRDRSINSILKRPLENLIPKYLTPVIOEFSHMLICRAGAPASHAGKVVQCK 300
DB 241 QVSPRDRSINSILKRPLENLIPKYLTPVIOEFSHMLICRAGAPASHAGKVVQCK 300
QY 301 IQKVRFGKCPPRSISVPIKRNAILHRNEWPPAGAKARSIMIERPKIAAVCGHYD 360
DB 301 IQKVRFGKCPPRSISVPIKRNAILHRNEWPPAGAKARSIMIERPKIAAVCGHYD 360
QY 361 YYAQLDMLRRRAHPSYHIPQENTGVEDYQETRHGSPSPQWPAEYLQKPEAQYK 420
DB 361 YYAQLDMLRRRAHPSYHIPQENTGVEDYQETRHGSPSPQWPAEYLQKPEAQYK 420
QY 421 VEKQLGLRPSAEPNYNQRELNSGEEPRFOELPFRKNEMKEQYWKOLEE---IROQY 477
DB 421 VEKQLGLRPSAEPNYNQRELNSGEEPRFOELPFRKNEMKEQYWKOLEE---IROQY 477
QY 478 L---NDKMEIRKXGPREP-EDIEKDLKQMLONTKESKNPEQKYAKKGVKPEINDKCI 533
DB 481 LPVHQDASE-----GEAPVQDIEKDLKQMLONTKESKNPEQKYAKKGVKPEINDKCI 535
QY 534 SDENILQEEAMDPNETLTFEDGKKEVECKVEHGDYTDKAFELKHCPEAAFTLT 591
DB 536 SDENILQEEAMDPNETLTFEDGKKEVECKVEHGDYTDKAFELKHCPEAGFSTQT 593

RESULT 5

ID AAU03545

XX AAU03545 standard; protein; 649 AA.

AC AAU03545;

XX 12-SEP-2001 (first entry)
XX
DE Human protein kinase #45.
XX
XX Human; protein kinase; PTK; STK; cancer; cardiovascular disease;
KW metabolic disorder; immune related disease; neurological disorder;
KW neurodegenerative disorder; inflammatory disorder; infectious disease;
KW reproductive disorder.
XX
OS Homo sapiens.
XX
PN WO200138503-A2.
XX
PD 31-MAY-2001.
XX
XX 22-NOV-2000; 2000WO-US032085.
PF
XX 24-NOV-1999; 99US-0167482P.
PR
XX (SUGEN-) SUGEN INC.
PA
XX Plowman GD, Whyte D, Manning G, Sudarsanam S, Martinez R;
PI Flanagan P, Clary D;
XX
DR WPI; 2001-343950/36.
DR N-PSDB; AAS06745.

XX Nucleic acids encoding human kinase polypeptides, useful for preventing
PT diagnosing and/or treating e.g. cancer, immune cardiovascular and
PT neuronal-associated diseases, and microbial infections.
PT
PS Claim 7; Fig 2; 433pp; English.

XX AAU03501-AAU03557 represent novel human protein kinases #1-57. The novel
CC protein kinases have been identified as members of the tyrosine or
CC serine/threonine kinase (PTK and STK) families. The polynucleotides
CC encoding protein kinases and the polypeptides may be used in the
CC prevention, diagnosis and treatment of diseases associated with
CC cancers (especially cancers of hematopoietic origin), cardiovascular
CC disease (e.g. atherosclerosis), metabolic disorders (e.g. diabetes),
CC immune related diseases (e.g. rheumatoid arthritis), neurological
CC disorders (e.g. schizophrenia), neurodegenerative disorders (e.g.
CC Parkinson's disease), inflammatory disorders (e.g. asthma), infectious
CC disease (e.g. HIV) and reproductive disorders (e.g. infertility).
CC Additionally, polynucleotides encoding protein kinases may be used for
CC gene therapy and as DNA probes in diagnostic assays. The protein kinase
CC polypeptides may be used as antigens in the production of antibodies
CC against the protein kinases and in assays to identify modulators of
CC protein kinase expression and activity
XX
SQ Sequence 649 AA;

Query Match 79.8%; Score 2715.5; DB 4; Length 649;

Best Local Similarity 88.4%; Pred. No. 2.7e-213;

Matches 526; Conservative 10; Mismatches 26; Indels 33; Gaps 5;

QY 1 MDKYDVIKAIQGGAGFKAYLAKGSDSKHCVIKENFEKMPIOKEASKKEVILEKWKH 60
DB 1 MDKYDVIKAIQGGAGFKAYLAKGSDSKHCVIKENFEKMPIOKEASKKEVILEKWKH 60
QY 61 PNIVAFNFSFQENGLFTWMEYCDGDLKMGINRQGVLFSEDIQILGFWQISLGLKHII 120
DB 61 PNIVAFNFSFQENGLFTWMEYCDGDLKMGINRQGVLFSEDIQILGFWQISLGLKHII 120
QY 121 DRKILHRDIKQNTIFLSKNGVAKLGFAGIARVLNNSMELARTCIGTPYLSPEICQNK 180
DB 121 DRKILHRDIKQNTIFLSKNGVAKLGFAGIARVLNNSMELARTCIGTPYLSPEICQNK 180
QY 181 YNNKTDIWSLGCVLVELCTLKHPEGNLQOLVLKI COAHFAPISPGFSRELHSLISOLF 240

Db 181 YNNKTDIWSLGCVLVELCTLKHPFEGNNLQQLVLTICQAHFAPISPGFSRELHSLISQLF 240
QY 241 QVSPDRPSINSILKRPFLNIPKYLTPDEVIOEE----FSHMLICRAGAPASRHAGKV 296
Db 241 QVSPDRPSINSILKRPFLNIPKYLTPDEVIOEE----FSHMLICRAGAPASRHAGKV 296
QY 297 QKCKIQKVRPRGKCPKPPRSISVPIKKNAILHRNWRPPAGAKARSIMIERPKIAAVCG 356
Db 289 ----LTRV-FGRCEPLHACIMVFN-----PSCFLQ-----IKMIERPKIAAVCG 331
QY 357 HYDYYAQLDMLRRRAHKPSYHPIPOENTGVEDYGOETRHGSPSQWPAEYLQRKFEAQ 416
Db 332 HYDYYAQLDMLRRRAHKPSYHPIPOENTGVEDYGOETRHGSPSQWPAEYLQRKFEAQ 391
QY 417 YLKVKEQLGLRPSAPPNYQQLRSNGEPRFOELPFRKNEMKEQYWKQLEIRQQ 476
Db 392 YLKVKEQLGLRPSAPPNYQQLRSNGEPRFOELPFRKNEMKEQYWKQLEIRQQ 451
QY 477 YLNDMKEIRKMGREPEDEKOLKQNLQNTKESKNPEQKYAKKGVKFEINLDCISDE 536
Db 452 YLNDMKEIRKMGREPEDEKOLKQNLQNTKESKNPEQKYAKKGVKFEINLDCISDE 511
QY 537 NILQEEBAMDIPNETLTFEDGMKFEYECVKEHGDYTDKAFKJLHCPPEAAFTLT 591
Db 512 NILQEEBAMDIPNETLTFEDGMKFEYECVKEHGDYTDKAFKJLHCPPEAGFTQT 566

RESULT 6

ID ABP71710
XX ABP71710 standard; protein; 631 AA.

AC ABP71710;

XX 17-APR-2003 (first entry)

XX Human NEK-like serine/threonine protein kinase # SEQ ID 2.

XX Human; NEK-like serine/threonine protein kinase; cytostatic; cardiant;
KW antiinflammatory; noctropic; neuroprotective; cancer; colon cancer;
KW cardiovascular disorder; diabetes; COPD; CNS disorder.

XX Homo sapiens.

XX WO200300903-A2.

XX 03-JAN-2003.

XX 24-JUN-2002; 2002WO-BF006948.

XX 25-JUN-2001; 2001US-0300068P.

PR 07-DEC-2001; 2001US-0336704P.

XX (FARB) BAYER AG.

XX Xiao Y;

XX WPI; 2003-184051/18.

DR N-PSDB; ABZ59716.

XX New polynucleotide encoding a NEK-like serine/threonine kinase

PT Polypeptide useful for treating diseases associated with kinase
PT dysfunction, e.g. cardiovascular disorders, cancer such as colon cancer,
PT diabetes and CNS disorders.

XX Claim 1 a ii; Fig 2; 149pp; English.

XX The invention relates to a newly isolated polynucleotide encoding an NEK-
CC like serine/threonine protein kinase. The activity of the polynucleotide
CC and polypeptide of the invention may be described as cytostatic,
CC cardiant, antiinflammatory, noctropic and neuroprotective. The expression
CC vector and reagent of the invention are useful for the preparation of a
CC medicament for modulating the activity of an NEK-like serine/threonine
CC kinase in a disease, such as cancer (e.g. colon cancer), cardiovascular

CC disorder, diabetes, COPD or CNS disorder. The polypeptides may also be
CC used to identify compounds which may act as activators or inhibitors at
CC the enzyme's active site, to raise specific antibodies which can block
CC the enzyme and effectively reduce its activity, as a bait protein in a
CC two-hybrid or three-hybrid assay to identify other proteins which bind to
CC or interact with the human NEK-like serine/threonine kinase polypeptide
CC and modulate its activity, and for the immunisation of mammals. The
CC current sequence represents the human NEK-like serine/threonine protein
CC kinase of the invention

XX SQ Sequence 631 AA;

Query Match 76.7%; Score 2610.5; DB 6; Length 631;

Best Local Similarity 85.8%; Pred. No. 1e-204;

Matches 512; Conservative 4; Mismatches 26; Indels 55; Gaps 5;

QY 1 MDKYDVVKAIGQAFGKAYLAKGSKHCVKEINFEKMPIQKEKSKKEVILEKMKH 60

Db 1 MDKYDVVKAIGQAFGKAYLAKGSKHCVKEINFEKMPIQKEKSKKEVILEKMKH 60

QY 61 PNIVAFNSFQENGRLEFIVMEYCDGDLMKRINRQGVLFSEDIQILGWFOISLGLKHH 120

Db 61 PNIVAFNSFQENGRLEFIVMEYCDGDLMKRINRQGVLFSEDIQILGWFOISLGLKHH 120

QY 121 DRKILHRDIAQNIIFLSKNGWAKLGDGFIARVLNNSMELARTCIGTPYVLSPEICQNK 180

Db 121 DRKILHRDIAQNIIFLSKNGWAKLGDGFIARVLNNSMELARTCIGTPYVLSPEICQNK 180

QY 181 YNNKTDIWSLGCVLVELCTLKHPPEGNNLQQLVLTICQAHFAPISPGFSRELHSLISQLF 240

Db 181 YNNKTDIWSLGCVLVELCTLKHPPEGNNLQQLVLTICQAHFAPISPGFSRELHSLISQLF 240

QY 241 QVSPDRPSINSILKRPFLNIPKYLTPDEVIOEESHMLICRAGAPASRHAGKVQKCK 300

Db 241 QVSPDRPSINSILKRPFLNIPKYLTPDEVIOEESHMLICRAGAPASRHAGKVQKCK 300

QY 301 IQKVRPGKCPSPRSISVP----IKRNAILHRNWRPPAGAKAR--SIKMIERPKIAAV 354

Db 298 ----RHGT---VRSLSRPWAULLRLCRLOPIRLISQAAVECLRLFOIKMIERPKIAAV 350

QY 355 CGHYDYYAQLDMLRRRAHKPSYHPIPOENTGVEDYGOETRHGSPSQWPAEYLQRKFEA 414

Db 351 CGHYDYYAQLDMLRRRAHKPSYHPIPOENTGVEDYGOETRHGSPSQWPAEYLQRKFEA 410

QY 415 QQYKLKVEKOLGLRPPSAEPNYNQRLSRNGEPRFQELPFRKNEMKEOEYKQLESEIR 474

Db 411 QQYKLKVEKOLGLRPPSAEPNYNQRLSRNGEPRFQELPFRKNEMKEOEYKQLESEIR 431

QY 475 QQYLDNMKEIRKMGREPEDEKDLKOMRLQNTKESKNPEQKYAKKGVKFEINLDCIS 534

Db 432 QQYHNDMKEIRKMGREPEDEKDLKOMRLQNTKESKNPEQKYAKKGVKFEINLDCIS 491

QY 535 DENILOBEEAMDIPNETLTFEDGMKFEYECVKEHGDYTDKAFKJLHCPPEAAFTLT 591

Db 492 DENILOBEEAMDIPNETLTFEDGMKFEYECVKEHGDYTDKAFKJLHCPPEAGFTQT 548

RESULT 7

ABU54635

ID ABU54635 standard; protein; 403 AA.

XX AC ABU54635;

XX 03-JUN-2003 (first entry)

XX Human NOVX polypeptide #94.

XX Human; NOVX; metabolic disorder; cardiomyopathy; diabetes; ASD;
KW hypertension; congenital heart defect; aortic stenosis; valve disease;
KW atrial septal defect; atrioventricular canal defect; ductus arteriosus;
KW pulmonary stenosis; subaortic stenosis; ventricular septal defect; VSD;
KW tuberculous sclerosis; scleroderma; atherosclerosis; infectious disease;
KW obesity; anorexia; neurodegenerative disorder; Alzheimer's disease;

KW Parkinson's disease; immune disorder; haematopoietic disorder;
 KW haemophilia; hypercoagulation; Crohn's disease; cancer.

OS Homo sapiens.

XX WO200281498-A2.

XX 17-OCT-2002.

XX 03-APR-2002; 2002WO-US010780.

XX 03-APR-2001; 2001US-0281086P.

XX 03-APR-2001; 2001US-0281136P.

XX 05-APR-2001; 2001US-0281863P.

XX 06-APR-2001; 2001US-0281906P.

XX 10-APR-2001; 2001US-0282930P.

XX 10-APR-2001; 2001US-0282934P.

XX 12-APR-2001; 2001US-0283512P.

XX 13-APR-2001; 2001US-0283710P.

XX 17-APR-2001; 2001US-0284234P.

XX 19-APR-2001; 2001US-0285325P.

XX 20-APR-2001; 2001US-0285381P.

XX 20-APR-2001; 2001US-0285609P.

XX 23-APR-2001; 2001US-0285748P.

XX 23-APR-2001; 2001US-0285890P.

XX 24-APR-2001; 2001US-0286068P.

XX 25-APR-2001; 2001US-0286292P.

XX 27-APR-2001; 2001US-0287213P.

XX 02-MAY-2001; 2001US-0288257P.

XX 29-MAY-2001; 2001US-0294164P.

XX 30-MAY-2001; 2001US-0294484P.

XX 18-JUN-2001; 2001US-0298952P.

XX 19-JUN-2001; 2001US-0299237P.

XX 19-JUN-2001; 2001US-0299276P.

XX 12-SEP-2001; 2001US-0324800P.

XX 25-SEP-2001; 2001US-0324802P.

XX 27-SEP-2001; 2001US-0325684P.

XX 17-OCT-2001; 2001US-0330143P.

XX 14-NOV-2001; 2001US-0332131P.

XX 14-NOV-2001; 2001US-0332240P.

XX 14-NOV-2001; 2001US-0332779P.

XX 21-NOV-2001; 2001US-0332115P.

XX 04-DEC-2001; 2001US-0337621P.

XX 03-JAN-2002; 2002US-0345783P.

XX 16-JAN-2002; 2002US-0350251P.

XX 02-APR-2002; 2002US-00114270.

(CURA-) CURAGEN CORP.

Guo X, Kekuda R, Miller CE, Malvankar UM, Spytek KA;

Patturajan M, Liu X, Gusev VT, Li L, Vernet CM, Zerhusen BD;

Gorman L, Shenoy SG, Pena CE, Smithson G, Burgess CE, Gerlach V;

Padigaru M, Shinkets RA, Gangoli EA, Taupier RJ, Casman SJ, Ji W;

Anderson DW, Leite MW, Rastelli L, Edinger SR, Stone DJ;

MacDougall JR, Rothenberg ME, Mazur A, Millet I, Peyman JA;

Ellerman K;

WPI; 2003-045858/04.

N-PSDB; ABX72263.

New isolated NOVX polypeptide useful for treating atherosclerosis,

metabolic disorders, diabetes, obesity, infectious disease, anorexia,

neurodegenerative disorders, Alzheimer's disease and cancer.

Claim 1; Page 292; 66pp; English.

The invention relates to human polypeptides, termed NOVX, and the

polynucleotides encoding them. The polypeptides and polynucleotides are

useful for diagnosing disease, and screening for potential therapeutic

agents. The sequences are useful for treating metabolic disorders,

cardiomyopathy, diabetes, hypertension, congenital heart defects, aortic

CC stenosis, atrial septal defect (ASD), atrioventricular canal defect,
 CC ductus arteriosus, pulmonary stenosis, subaortic stenosis, ventricular
 CC septal defect (VSD), valve diseases, tuberosus sclerosis, scleroderma,
 CC atherosclerosis, obesity, infectious disease, anorexia, neurodegenerative
 CC disorders, Alzheimer's disease, Parkinson's disease, immune disorders,
 CC haematopoietic disorders, haemophilia, hypercoagulation, Crohn's disease
 CC and cancer. Sequences ABU54542-ABU54647 represent human NOVX polypeptides
 CC of the invention

XX SQ Sequence 403 AA;

Query Match 61.9%; Score 2107.5; DB 6; Length 403;

Best Local Similarity 98.8%; Fred. No. 9e-164;

Matches 398; Conservative 3; Mismatches 1; Indels 1; Gaps 1;

QY 1 MDKYDVIRKAIQGAGFKAYLAKGSDSKHCYKVIKEINFEKMPIQEKEASKKEVILKKMH 60

Db 1 MDKYDVIRKAIQGAGFKAYLAKGSDSKHCYKVIKEINFEKMPIQEKEASKKEVILKKMH 60

QY 61 PNIVAFNSFQENGRLFIWMEYCDGGDLMKGINRQGVLPFSEDQILGFVQISLGLKH 120

Db 61 PNIVAFNSFQENGRLFIWMEYCDGGDLMKGINRQGVLPFSEDQILGFVQISLGLKH 120

QY 121 DRKILHRDIKQNIPLSKNGVMVAKLDFGIARVLNNSMELARTCIGTPYILSPICQNK 180

Db 121 DRKILHRDIKQNIPLSKNGVMVAKLDFGIARVLNNSMELARTCIGTPYILSPICQNK 180

QY 181 YNNKTDIWSLGLCVLYELCTLKHPFEGNNLQQLVLIKCOAHFAPISPGFSRHLHSLIQL 240

Db 181 YNNKTDIWSLGLCVLYELCTLKHPFEGNNLQQLVLIKCOAHFAPISPGFSRHLHSLIQL 240

QY 241 QVSPDRPSINSILKRPENLIPKYLTPV-IOEEFESHMLICRAGAPASHAGKVQKC 299

Db 241 QVSPDRPSINSILKRPENLIPKYLTPV-IOEEFESHMLICRAGAPASHAGKVQKC 300

QY 300 KIQKVRFRGKCPPRSISVPIKRNAILHNEWPPAGAKARSIMIERPKIAAVCGHYD 359

Db 301 KIQKVRFRGKCPPRSISVPIKRNAILHNEWPPAGAKARSIMIERPKIAAVCGHYD 360

QY 360 YYYAQLDMLRRRAKPSYHPIQENTGVEDYGOETRHGSPSQ 402

Db 361 YYYAQLDMLRRRAKPSYHPIQENTGVEDYGOETRHGSPSQ 403

RESULT 8

ABU54634

ID ABU54634 standard; protein; 399 AA.

XX AC ABU54634;

XX XX 03-JUN-2003 (first entry)

XX DE Human NOVX polypeptide #93.

XX KW Human; NOVX; metabolic disorder; cardiomyopathy; diabetes; ASD;

XX KW hypertension; congenital heart defect; aortic stenosis; valve disease;

XX KW atrial septal defect; atrioventricular canal defect; ductus arteriosus;

XX KW pulmonary stenosis; subaortic stenosis; ventricular septal defect; VSD;

XX KW tuberosus sclerosis; scleroderma; atherosclerosis; infectious disease;

XX KW obesity; anorexia; neurodegenerative disorder; Alzheimer's disease;

XX KW Parkinson's disease; immune disorder; haematopoietic disorder;

XX KW haemophilia; hypercoagulation; Crohn's disease; cancer.

XX OS Homo sapiens.

XX XX WO200281498-A2.

XX PD 17-OCT-2002.

XX PF 03-APR-2002; 2002WO-US010780.

XX PR 03-APR-2001; 2001US-0281086P.

XX PR 03-APR-2001; 2001US-0281136P.

PR 05-APR-2001; 2001US-0281863P.
 PR 05-APR-2001; 2001US-0281906P.
 PR 06-APR-2001; 2001US-0282020P.
 PR 10-APR-2001; 2001US-0282930P.
 PR 10-APR-2001; 2001US-0282934P.
 PR 12-APR-2001; 2001US-0283512P.
 PR 13-APR-2001; 2001US-0283710P.
 PR 17-APR-2001; 2001US-0284234P.
 PR 19-APR-2001; 2001US-0285325P.
 PR 20-APR-2001; 2001US-0285381P.
 PR 20-APR-2001; 2001US-0285609P.
 PR 23-APR-2001; 2001US-0285748P.
 PR 23-APR-2001; 2001US-0285890P.
 PR 24-APR-2001; 2001US-0286068P.
 PR 25-APR-2001; 2001US-0286292P.
 PR 27-APR-2001; 2001US-0287213P.
 PR 02-MAY-2001; 2001US-0288257P.
 PR 25-MAY-2001; 2001US-0294164P.
 PR 30-MAY-2001; 2001US-0294484P.
 PR 18-JUN-2001; 2001US-0298952P.
 PR 19-JUN-2001; 2001US-0299237P.
 PR 19-JUN-2001; 2001US-0299276P.
 PR 12-SEP-2001; 2001US-0318750P.
 PR 25-SEP-2001; 2001US-0324800P.
 PR 25-SEP-2001; 2001US-0324802P.
 PR 27-SEP-2001; 2001US-0325684P.
 PR 17-OCT-2001; 2001US-0330143P.
 PR 14-NOV-2001; 2001US-0332131P.
 PR 14-NOV-2001; 2001US-0332240P.
 PR 14-NOV-2001; 2001US-0332779P.
 PR 21-NOV-2001; 2001US-0332115P.
 PR 04-DEC-2001; 2001US-0337621P.
 PR 03-JAN-2002; 2002US-0345783P.
 PR 16-JAN-2002; 2002US-0350251P.
 PR 02-APR-2002; 2002US-00114270.
 XX
 PA (CURA-) CURAGEN CORP.
 XX
 PI Guo X, Kekuda R, Miller CE, Malyankar UM, Spytek KA;
 PI Patturajan M, Liu X, Gusev VT, Li L, Vernet CAM, Zethusen BD;
 PI Gorman L, Shenoy SG, Pena CE, Smithson G, Burgess CE, Gerlach V;
 PI Padigaru M, Shinkets RA, Gangolli EA, Taupier RJ, Casman SJ, Ji W;
 PI Anderson DW, Leite MW, Rastelli L, Edinger SR, Stone DJ;
 PI MacDougall JR, Rothenberg ME, Mazur A, Millet I, Peyman JA;
 PI Ellerman K;
 XX
 XX WPI; 2003-046858/04.
 DR N-PSDB; ABX72262.
 XX
 PT New isolated NOVX polypeptide useful for treating atherosclerosis,
 PT metabolic disorders, diabetes, obesity, infectious disease, anorexia,
 PT neurodegenerative disorders, Alzheimer's disease and cancer.
 XX
 PS Claim 1; Page 291; 666pp; English.
 XX
 CC The invention relates to human polypeptides, termed NOVX, and the
 CC polynucleotides encoding them. The polypeptides and polynucleotides are
 CC useful for diagnosing disease, and screening for potential therapeutic
 CC agents. The sequences are useful for treating metabolic disorders,
 CC cardiomyopathy, diabetes, hypertension, congenital heart defects, aortic
 CC stenosis, atrial septal defect (ASD), atriocentric stenosis, ventricular
 CC ductus arteriosus, pulmonary stenosis, subaortic stenosis, scleroderma,
 CC atherosclerosis, obesity, infectious disease, Parkinson's disease, immune disorders,
 CC disorders, Alzheimer's disease, Crohn's disease, Crohn's disease
 CC haematopoietic disorders, haemophilia, hypercoagulation, Crohn's disease
 CC and cancer. Sequences ABUS4542-ABUS4647 represent human NOVX polypeptides
 CC of the invention
 XX
 SQ Sequence 399 AA;

Query Match 61.8%; Score 2102.5; DB 6; Length 399;
 Best Local Similarity 99.0%; Pred. No. 2.3e-163;

Matches 398; Conservative 1; Mismatches 0; Indels 3; Gaps 1;
 QY 1 MDKYDVIAICQAGAFKAYLAKGKSDSKHCVIKINFERKPIQKEASKKEVILLERKWKH 60
 DB 1 MDKYDVIAICQAGAFKAYLAKGKSDSKHCVIKINFERKPIQKEASKKEVILLERKWKH 57
 QY 61 PNIVAFNFSQENGRLEFIVMEYCDGDLMKRINRQGVLFSEDOILGWFQVQISLGLKHIIH 120
 DB 58 PNIVAFNFSQENGRLEFIVMEYCDGDLMKRINRQGVLFSEDOILGWFQVQISLGLKHIIH 117
 QY 121 DRKILHRDIKAQNTIFLSKNGWAKLGFGIARVLNNSMELARTICIGTPYYLSPEICQK 180
 DB 118 DRKILHRDIKAQNTIFLSKNGWAKLGFGIARVLNNSMELARTICIGTPYYLSPEICQK 177
 QY 181 YNNKTIDWSLGCVLVELCTLKHPPEGNLQOLVKICQAHFAPISPGFSRELHSLISQLF 240
 DB 178 YNNKTIDWSLGCVLVELCTLKHPPEGNLQOLVKICQAHFAPISPGFSRELHSLISQLF 237
 QY 241 QVSPDRPSINSILKRPFLLENLIPKYLTPVIOEFSHMLICRAGAPASHAGKVKQCK 300
 DB 238 QVSPDRPSINSILKRPFLLENLIPKYLTPVIOEFSHMLICRAGAPASHAGKVKQCK 297
 QY 301 IQKVFQKCPKPSRISVPIKRNAILHRNWRPPAGAKAKARSIMTERPKIAAVCGHYDY 360
 DB 298 IQKVFQKCPKPSRISVPIKRNAILHRNWRPPAGAKAKARSIMTERPKIAAVCGHYDY 357
 QY 361 YYAQLDMLRRRAHKPSYHPPIQENTGVEDYQETRHGSPSQ 402
 DB 358 YYAQLDMLRRRAHKPSYHPPIQENTGVEDYQETRHGSPSQ 399
 RESULT 9
 ADF44515
 ID ADF44515 standard; protein; 614 AA.
 XX
 AC ADF44515;
 XX
 DT 12-FEB-2004 (first entry)
 DE Mouse kinase protein SEQ ID NO:33.
 XX
 KW cytostatic; nootropic; neuroprotective; antidiabetic; screening;
 KW regulation; drug development; protein-associated disease; cancer;
 KW dementia; diabetes; kinase; enzyme; mouse.
 XX
 OS Mus musculus.
 XX
 PN WO2003084992-A1.
 XX
 PD 16-OCT-2003.
 XX
 PF 04-APR-2003; 2003WO-JP004330.
 XX
 PR 05-APR-2002; 2002JP-00103396.
 PR 23-APR-2002; 2002JP-00120904.
 PR 02-MAY-2002; 2002JP-00130601.
 PR 04-DEC-2002; 2002JP-00352520.
 XX
 PA (RIKE) RIKEN KK.
 PA (DNAP-) DNAPFORM KK.
 PA (MITU) MITSUBISHI CHEM CORP.
 XX
 PI Hayashizaki Y, Kamiya M, Kubodera H, Watanabe W;
 DR WPI; 2003-833568/77.
 DR N-PSDB; ADF44489.
 XX
 PT Proteins and encoded DNAs with kinase activity, useful in screening
 PT substances for regulating such activity and in developing drugs for the
 PT protein-associated diseases e.g. cancer, dementia and diabetes.
 XX
 PS Claim 1; SEQ ID NO 33; 342pp; Japanese.
 XX

The present invention describes a protein: (a) containing any of the amino acid sequences of ADF44509 to ADF44534 or ADF44544; or (b) based on any of the sequences in (a) but with some amino acids deleted, substituted and/or added and having kinase activity. Also described: (1) a DNA encoding any of the proteins; (2) a full-length cDNA encoding the protein; (3) a DNA which is: (a) a DNA containing any of the base sequences in ADF44483 to ADF44508 or ADF44543; (b) a DNA derived from any of the sequences in (a) but with some bases deleted, substituted and/or added and encoding a protein with kinase activity; or (c) a DNA hybridizable with any of the sequences in (a) or their complementary strands under stringent conditions and encoding a protein with kinase activity; (4) a recombinant vector containing the DNA; (5) a cell transfected with the DNA or recombinant vector, or an individual produced from the cell; (6) recombinant proteins produced by such cells; (7) an oligonucleotide containing 5-10 consecutive bases in any of the base sequences, its sense oligonucleotide, an antisense oligonucleotide with a complementary strand of such sense oligonucleotide, or an oligonucleotide derivative of the (anti-)sense oligonucleotide; (8) an antibody specifically binding to the protein, or its partial fragment; (9) a method for screening substances for regulating activity of the protein by contacting a test substance with such protein before measuring changes in the protein activity due to the test substance; (10) a method for screening substances regulating expression of the DNA by contacting a test substance with cells transfected with the gene and detecting changes in expression level of the DNA in such cells; (11) recordable media for reading in a computer with information on the amino acid sequences of the proteins, and/or base sequences of the DNAs stored; and (12) a support for binding with any of the proteins and/or DNAs. The proteins and their encoded DNAs have cytostatic, neurotropic, neuroprotective and antidiabetic activities. They can be used in screening substances for regulating such activity and in developing drugs for the protein-associated diseases e.g. cancer, dementia and diabetes. The present sequence is used in the exemplification of the present invention.

Sequence 614 AA;

Query Match 48.2%; Score 1640.5; DB 7; Length 614;
 Best Local Similarity 54.8%; Pred. No. 2.9e-125;
 Matches 340; Conservative 49; Mismatches 72; Indels 159; Gaps 9;

QY 1 MDKTDVKAICQGAAGKAYLAKGSDSKHCVIKENFERKPIQEKASKEVILLKMKH 60
 DB 1 MDNFHLIKIIGEGTFGKYLAKDSSESHCVIKESLTK----EKEASKEVILLARMEH 56

QY 61 PNIVAFNFSFQENGLFTVMEYCDGDLMKRINRQGVLFSEDOILGWVQVQISLGLKHH 120
 DB 57 PNIVTFSSFQENGLFTVMEYCDGDLQRIQQRGVMEFSEDOILGWVQVQISLGLKHH 116

QY 121 DRKILHRDIKAQNTFLSKNGWAKLGFDTIARVLNNSMELARTICGTPYVLSPEICQKP 180
 DB 117 DRKILHRDIKSNITFLSKNGWAKLGFDTIARVLNNSMELARTICGTPYVLSPEICQNP 176

QY 181 YNNKTDIWSLGCVLVELCTLKHPEGNLQLOVLKICQAHFAPISPGFSRELHSLISOLF 240
 DB 177 YNNKTDIWSLGCVLVELCTLKHPEGNLQLOVLKICQAHFAPISPGFSRELHSLISOLF 236

QY 241 QVSPDRPFSINSILKRPFLNLIPIKLTPTVQIEFSLHMLICRAGAPASHACKVQKCK 300
 DB 237 RVSPQDRPFSVTLKRPFLNLIPIKLTPTVQIEFSLHMLICRAGAPASHACKVQKCK 267

QY 301 IQKVRFRGCKPFRSRIQVPIKRNAILHRNWRPAGAKARSIKWIERPKIAVCGHYD 360
 DB 268 -----CSR----- 271

QY 361 YYAQIDMLRRRAHKPSVHIPQENTGVEDYQETRHGESP-----SOWPBYLQKFEAAQ 416
 DB 272 -----IQSHAH-----VENMAI-----GPTACWRVSPWSAAYLQKFEAAQ 307

QY 417 YKLKVEKQLGRLPSAENYVQRQLNSGEPFQELPRKQEMKQEVYKQLEIRIQ 476
 DB 308 YKLKVEKQLGRLPSVPEHPNEGKQLQSHWEETKQELQVRKQKMKQDEYKQLEIRIQ 367

QY 477 YLNDMKEIRKQNGREPDIENDLQKMRILQNTKESKNPEQYKAKKGVKFEINLDKCISE 536

DB 368 YHNDMKEIRKQNGREPDIENDLQKMRILQNTKESKNPEQYKAKKGVKFEINLDKCISE 400
 QY 537 NILQEEAMDIPINETLTFEDGMKFEKVEYCKEHEGDTYDKAFKELHCPD-----AAF 587
 DB 401 DTVQENEAVDKLNATLSFEDGTQFQHRCKEHEGDTYDKAFKELHCPD-----AAF 460

QY 588 TELTWLSFLFLEYSLPHLL 607
 DB 461 NRQW-----DAGAPHTLL 474

RESULT 10
 ADF44528
 ID ADF44528 standard; protein; 336 AA.
 XX ADF44528;
 AC ADF44528;
 DT 12-FEB-2004 (first entry)
 XX Mouse kinase protein SEQ ID NO:46.
 DE
 XX Cytostatic; neurotropic; neuroprotective; antidiabetic; screening;
 KW regulation; drug development; protein-associated disease; cancer;
 KW dementia; diabetes; kinase; enzyme; mouse.
 OS Mus musculus.
 XX
 XX WO2003084992-A1.
 FN
 PD 16-OCT-2003.
 XX
 XX 04-APR-2003; 2003WO-JP004330.
 XX
 XX 05-APR-2002; 2002JP-00103396.
 PR 23-APR-2002; 2002JP-00120904.
 PR 02-MAY-2002; 2002JP-00130601.
 PR 04-DEC-2002; 2002JP-00352520.
 XX
 XX (RIKE) RIKEN KK.
 PA (DNAP-) DNAFORM KK.
 PA (MITU) MITSUBISHI CHEM CORP.
 XX
 PI Hayashizaki Y, Kamiya M, Kubodera H, Watanabe W;
 XX
 XX WPI: 2003-833568/77.
 DR N-PSDB; ADF44502.
 XX
 XX Proteins and encoded DNAs with kinase activity, useful in screening
 PT substances for regulating such activity and in developing drugs for the
 PT protein-associated diseases e.g. cancer, dementia and diabetes.
 XX
 PS Claim 1; SEQ ID NO 46; 342pp; Japanese.
 CC
 CC The present invention describes a protein: (a) containing any of the
 CC amino acid sequences of ADF44509 to ADF44534 or ADF44544; or (b) based on
 CC any of the sequences in (a) but with some amino acids deleted,
 CC substituted and/or added and having kinase activity. Also described: (1)
 CC a DNA encoding any of the proteins; (2) a full-length cDNA encoding the
 CC protein; (3) a DNA which is: (a) a DNA containing any of the base
 CC sequences in ADF44483 to ADF44508 or ADF44543; (b) a DNA derived from any
 CC of the sequences in (a) but with some bases deleted, substituted and/or
 CC added and encoding a protein with kinase activity; or (c) a DNA
 CC hybridizable with any of the sequences in (a) or their complementary
 CC strands under stringent conditions and encoding a protein with kinase
 CC activity; (4) a recombinant vector containing the DNA; (5) a cell
 CC transfected with the DNA or recombinant vector, or an individual produced
 CC from the cell; (6) recombinant proteins produced by such cells; (7) an
 CC oligonucleotide containing 5-10 consecutive bases in any of the base
 CC sequences, its sense oligonucleotide, an antisense oligonucleotide with a
 CC complementary strand of such sense oligonucleotide, or an oligonucleotide
 CC derivative of the (anti-)sense oligonucleotide; (8) an antibody
 CC specifically binding to the protein, or its partial fragment; (9) a

PF	21-JUN-2002; 2002WO-BP006879.	
XX		
PR	25-JUN-2001; 2001US-0300071P.	
PR	16-NOV-2001; 2001US-0331447P.	
PR	07-DEC-2001; 2001US-0336693P.	
XX		
PA	(FARB) BAYER AG.	
XX		
PI	Xiao Y;	
XX		
DR	WPI; 2003-201424/19.	
DR	N-PSDB; AAL51590.	
XX		
PT	New serine/threonine protein kinase NEK1 gene and protein, useful for	
PT	identifying modulators of serine/threonine protein kinase NEK1 activity,	
PT	and in gene therapy for treating cancer, diabetes, heart failure or	
PT	Alzheimer's disease.	
XX		
PS	Claim 1; Fig 2; 156pp; English.	
XX		
CC	The invention comprises the amino acid and coding sequence of the human	
CC	serine/threonine protein kinase NEK1. The DNA and protein sequences of	
CC	the invention are useful for modulating the activity of serine/threonine	
CC	kinase NEK1 in a disease, such as: cancer (particularly colon cancer);	
CC	cardiovascular disorders; central nervous system (CNS) disorders;	
CC	diabetes; and chronic obstructive pulmonary disease. In particular the	
CC	DNA and protein sequences of the invention are useful for treating:	
CC	congestive heart failure; myocardial infarction; ischaemic heart disease;	
CC	arrhythmia; hypertensive; Alzheimer's disease; Parkinson's disease; and	
CC	peripheral or chronic pain. The present amino acid sequence represents	
CC	the human serine/threonine protein kinase NEK1 of the invention	
XX		
SQ	Sequence 1242 AA;	
Query Match		
Best Local Similarity 31.9%; Score 1086; DB 6; Length 1242;		
Matches 278; Conservative 122; Mismatches 197; Indels 254; Gaps 21;		
QY	1 MDKYVIVKAIAGGAFKAYLAKGSDSKHCIVIKINFEKMPITQEKESKVEKLVILEKMKH 60	
DB	1 MEKYVYVQIKIGSGFKAILVKSTEDGRQYVKEINISMSKKEESREVRVAVLANMKH 60	
QY	61 PHIVAFFNSFQNGRLFIWMEVCDGDLWRINRQGVLFSDQILGWVQISLGLKHH 120	
DB	61 PHIVQYRESFEENGSLYIMVDYCEGDLFRKINAQKGVLFQBDQILDWFVQICLAKHHV 120	
QY	121 DRKILHRDIKAQNIIFLSKNGMAKLDGFIARVLANSMELARTCIGTPYVLSPEICNKP 180	
DB	121 DRKILHRDIKSNIFLTKDGTV-QLQDGFARVLANSTVELARTCIGTPYVLSPEICNKP 179	
QY	181 YNNKTDIWSLGVLYELCTLKHPPFGNNLQQLVKICOAHFAPISPGFSRELHLSISOLF 240	
DB	180 YNNKSDIWAIGCVLYELCTLKHAFAGSKNVLVKIISGSPFPVSLHYSDLSIVLSOLF 239	
QY	241 QVSPDRPSINSILKRPFLNLIKVLTPVETIQEESFHLICRAGA---PASRHAG--- 293	
DB	240 KKNPDRPSVNSILEKGFIAKRIEFLSPQIAEFCUKTFSKFSQSQIPAKRPASQONS 299	
QY	294 -KVVQKCKIQKVRFGKCPFRSRISVPI-----KRNAILHENEWRPP----- 334	
DB	300 ISVMPAQKITK-----PAAKYGIPLAYKYGDKLHKPLQKQKHAHQTEKRYNT 351	
QY	335 -----AGAKARSIKWIER----- 348	
DB	352 GEERKIKSEAAKRRLEFIEKEKKQDQIISLMKAEQMKRQEKERLERINRAREQGRN 411	
QY	349 -----PKIAAVCGHYDYYAQLDMLRRRAHKPSVHPPIQENTG 386	
DB	412 VLSAGSGEVKAPFLGSGGTINAPSSFSRQYEHVHAIFDQMQQQRADN-----EAKWK 466	
QY	387 VEDYG-----QETHRGSSPSQWPAEYLQKRPQAYQYKLVKEQKGL-----RPS 430	
DB	467 REIYGRGLPERQKGLAVERAKQVEFLQKREANQKARAEHGMILQNLAAAMYGGRRPS 526	
QY	431 SABENYNQROELRSNGEPRFQELPPFRKNEMKSOEYKWLQEEIRQQYVNDMKIRKQNGR 490	
DB	527 S-----SRGKPRNKE-----EEVYLARLQIRLQNFNERQQIKAKLRG 565	
QY	491 EPEDI-----EKDLKQMBLQNTK-----ESKNPE-----OKY 517	
DB	566 EKKEANHSEGGEGSEADMRKIKIESLKAHANARAAVLKEQLERKKEAYERKKVWEEH 625	
QY	518 KAKGVKVF-----EINLDKICISDENILOEE--- 542	
DB	626 LVAKGVKSDVSPPLGQHETGSGSPSKQMRSVISVTSALKVEGVDSLSLTDRETSEMQX 685	
QY	543 -----BAMDIPNETLTFEDGMKFE-----YECVKEHGDYTDKAFELKHP 583	
DB	686 TNNAISSKRILRLRNENLKAQDEKQKQNLSTFTFINVEDAKHEKEKSVSSDRKKW- 744	
QY	584 EAAFTLTWLSFLFLEYSFLPHFLLEKSPFSRHILIEDLL-----CANDCSLK 629	
DB	745 EAGQQLVPLDELTLDTSF-----STTERHTVGEVILKGPNGSPRAWGKSPDTSVLK 797	
QY	630 DWSEKEMELRT 640	
DB	798 ILGEAELOLOT 808	
RESULT 14		
AAO26613		
ID	AAO26613 standard; protein; 774 AA.	
XX	AC	
XX	AAO26613;	
DT	28-MAR-2003 (first entry)	
DE	Serine/threonine protein kinase-like enzyme related mouse protein.	
XX		
KW	Cytostatic; antiparkinsonian; rootropic; neuroprotective; cardiatic;	
KW	hypotensive; antiarrhythmic; antianal; analgesic; antiinflammatory;	
KW	antidiabetic; tranquilizer; antianemic; antidepressant; uropathic;	
KW	anorectic; serine/threonine protein kinase; enzyme; cancer;	
KW	central nervous system disorder; cardiovascular disorder; diabetes;	
KW	chronic obstructive pulmonary disease; obesity; genito-urinary system;	
XX	pain; genetic testing; mouse.	
OS	Mus musculus.	
XX		
PN	WO200299096-A1.	
XX		
PD	12-DEC-2002.	
XX		
PF	06-JUN-2002; 2002WO-BP006203.	
XX		
PR	07-JUN-2001; 2001US-0296164P.	
PR	19-SEP-2001; 2001US-0323100P.	
PR	25-OCT-2001; 2001US-0330578P.	
PR	17-JAN-2002; 2002US-0348601P.	
XX		
PA	(FARB) BAYER AG.	
XX		
PI	Smolyar A, Horner EJ, Thelwell C;	
XX		
DR	WPI; 2003-140620/13.	
XX		
PT	New human serine/threonine protein kinase-like enzyme polypeptide and	
PT	polynucleotide, useful for regulating the activity of the protein kinase-	
PT	like enzyme to prevent, treat or ameliorate diabetes, cancer or obesity.	
XX		
PS	Disclosure; Fig 3; 157pp; English.	
XX		
CC	The invention relates to an isolated polynucleotide comprising: a	
CC	polynucleotide encoding a serine/threonine protein kinase-like enzyme	
CC	polypeptide; a sequence of 840, 990, 1120, 2142, or 1383 base pairs fully	
CC	defined in the specification; a polynucleotide that hybridises to the	

CC serine/threonine protein kinase-like enzyme DNA or a fragment thereof; or
CC a degenerate analogue of the said polynucleotides. The serine/threonine
CC protein kinase-like enzyme polypeptide and polynucleotide are useful in
CC preventing, ameliorating, or treating diseases associated with serine/
CC threonine protein kinase-like enzyme dysfunction such as cancer, central
CC nervous system disorders, cardiovascular disorders, chronic obstructive
CC pulmonary disease, diabetes, obesity, or disorders of the genito-urinary
CC system. These can also be used to treat pain associated with the above
CC disorders. The serine/threonine protein kinase-like enzyme polypeptide is
CC also useful in diagnostic assays or in genetic testing. The expression
CC vector or the reagent is useful in preparing a medicament for modulating
CC the activity of a serine/threonine protein kinase-like enzyme in a
CC disease, e.g. cardiovascular disorder, obesity, a disorder of the genito-
CC urinary system, a central nervous system disorder, diabetes, cancer, or
CC chronic obstructive pulmonary disease. The methods are useful in
CC producing and detecting the polynucleotide and polypeptide and in
CC screening for agents that modulate the activity of the serine/threonine
CC protein kinase-like enzyme polypeptide. This sequence represents a mouse
CC protein relating to the serine/threonine protein kinase-like enzyme
CC of the invention
XX
SQ Sequence 774 AA;

Query Match 31.7%; Score 1078; DB 6; Length 774;
Best Local Similarity 36.4%; Pred. No. 4.7e-79;
Matches 260; Conservative 107; Mismatches 171; Indels 176; Gaps 21;
QY 1 MDKYLVIKALGOGAFKAYLAKGSDSKHCVKEINFEKMPIOEKASKKEVILLKMKH 60
DB 1 MEKVRLQKIGEGFGKAVLVKSTEDGRHYVKEINISRMDSKQESREAVLANMKH 60
QY 61 PNIVAFNSFOENGRLEFVMEYCDGDLMKRINRQGVLFSEQILGWFWQISGLKHH 120
DB 61 PNIVQYKESFEENGLYIVMDYCEGGDLFKRINAQKALFQEDQILDWFVQICLAKH 120
QY 121 DRKILHRDIKAQNTIFLKGWAKLGFGIARVLNLSMELARTICIGTPYVLSPEICNKP 180
DB 121 DRKILHRDIKSNIFLTKGTV-QLGDFGIARVLNLSMELARTICIGTPYVLSPEICNKP 179
QY 181 YNNKTDIWSLGCVLVELCTLKHPEGNLQOLVLIKCOAHFAPISPGFSELHSLISQLF 240
DB 180 YNNKSDIHWALGCVLYELCTLKHPEAGNKNLVKIIISGSPFPVSHYSYDLRSLISQLF 239
QY 241 QVSPDRPSINSILKRFLENLPKYLTPVTEQESRSHMLICRAG----- 285
DB 240 KRNPRDRPSVNSILKRFLENLPKYLTPVTEQESRSHMLICRAG----- 285
QY 286 -----APASRRHAGVVKCKIQKVRFGKCPSPRSIS--VPKRNAILHRNEWRP 333
DB 300 SSFVPAQKITPAKYGVELTYKYGDKLLEKPPPKHQAHQIPVK--MNSGEERK 356
QY 334 PAG--AKAKSIKMIER----- 348
DB 357 KMSEAAKRRLEFIEKEKQKQDQIRFLKABQMKRQKQKLERINRAREQGNVRLRAGG 416
QY 349 -----PKIAVCGHYDYVAQLDMLRR-RAH-----KPSYHP--IPQE 383
DB 417 SGEVYKASFFGIGGAVSPSPCGQYEHYHAI PQMQRLAEADNEARWKGKGIYGRWLPER 476
QY 384 NTG---VEDYQETRHGSPSQWPAEYLRQKFRQYKLVKVE-----KOLGURPS 431
DB 477 QKQHLAVERANQ-----VEEFLQRKRAMQKARAEHVVYVYLAQLRQLRLQ--- 522
QY 432 AEPYNVORQELR-----SNGEPRFQBLPRKNEKMEQEWYKQLEBEIRQOYLN 479
DB 523 ---NFNERQQIKALRGKENKADTKQCEAT-BETDVR-----LKWQSLKATNARAVLK 575
QY 480 DMKEIRKQMGREPEDIKDLQKQMLQNTKESKNP-----BQYKAKKGVKF--- 525
DB 576 EQLERKKEAYERE--KKWEEHLVARVKSVDVPLPLELLETGSPSPSKQVQKVPISVTS 633
QY 526 --EINLKKCISDENILQFEEDAMDIPNETLT-----FEDGMKFYEYECVKEH 569

DB 634 LKEVLDGSLTD---TQEEEMKSNAGSISKREILRLNENLKAQDEKEKQH 683
RESULT 15
ABP71711
ID ABP71711 standard; protein; 774 AA.
XX AC ABP71711;
XX DT 17-APR-2003 (first entry)
XX DE NEK-like serine/threonine protein kinase NEK1_mouse # SEQ ID 3.
XX KW NEK-like serine/threonine protein kinase; cytostatic; cardiant;
XX KW antiinflammatory; nootropic; neuroprotective; cancer; colon cancer;
XX KW cardiovascular disorder; diabetes; COPD; CNS disorder; mouse; rat.
XX OS Mus musculus.
XX OS Rattus norvegicus.
XX PN WO200300903-A2.
XX XX 03-JAN-2003.
XX XX 24-JUN-2002; 2002WO-EP006948.
XX PR 25-JUN-2001; 2001US-0300068P.
XX PR 07-DEC-2001; 2001US-0336704P.
XX PA (FARB) BAYER AG.
XX PI Xiao Y;
XX DR WPI; 2003-184051/18.
XX PT New polynucleotide encoding a NEK-like serine/threonine kinase
XX PT polypeptide useful for treating diseases associated with kinase
XX PT dysfunction, e.g. cardiovascular disorders, cancer such as colon cancer,
XX PT diabetes and CNS disorders.
XX PS Disclosure; Fig 3; 149pp; English.
XX CC The invention relates to a newly isolated polynucleotide encoding an NEK-
XX CC like serine/threonine protein kinase. The activity of the polynucleotide
XX CC and polypeptide of the invention may be described as cytostatic,
XX CC cardiant, antiinflammatory, nootropic and neuroprotective. The expression
XX CC vector and reagent of the invention are useful for the preparation of a
XX CC medicament for modulating the activity of an NEK-like serine/threonine
XX CC kinase in a disease, such as cancer (e.g. colon cancer), cardiovascular
XX CC disorder, diabetes, COPD or CNS disorder. The polypeptides may also be
XX CC used to identify compounds which may act as activators or inhibitors at
XX CC the enzyme's active site, to raise specific antibodies which can block
XX CC the enzyme and effectively reduce its activity, as a bait protein in a
XX CC two-hybrid or three-hybrid assay to identify other proteins which bind to
XX CC and interact with the human NEK-like serine/threonine kinase polypeptide
XX CC and modulate its activity, and for the immunisation of mammals. The
XX CC current sequence represents an NEK-like serine/threonine protein kinase
XX CC designated NEK1_mouse
XX SQ Sequence 774 AA;

Query Match 31.7%; Score 1078; DB 6; Length 774;
Best Local Similarity 36.4%; Pred. No. 4.7e-79;
Matches 260; Conservative 107; Mismatches 171; Indels 176; Gaps 21;
QY 1 MDKYLVIKALGOGAFKAYLAKGSDSKHCVKEINFEKMPIOEKASKKEVILLKMKH 60
DB 1 MEKVRLQKIGEGFGKAVLVKSTEDGRHYVKEINISRMDSKQESREAVLANMKH 60
QY 61 PNIVAFNSFOENGRLEFVMEYCDGDLMKRINRQGVLFSEQILGWFWQISGLKHH 120
DB 61 PNIVQYKESFEENGLYIVMDYCEGGDLFKRINAQKALFQEDQILDWFVQICLAKH 120

Search completed: November 12, 2004, 13:03:04
Job time : 162 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 7, 2004, 20:25:58 ; Search time 9743 Seconds
(without alignments)
11503.293 Million cell updates/sec

Title: US-10-730-010-1

Perfect score: 2370

Sequence: 1 cattggagaccatggataag.....ctgtcatgccgttacgagcg 2370

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 4526729 seqs, 23644849745 residues

Total number of hits satisfying chosen parameters: 9053458

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

GenEmbl.*

1: gb_ba.*

2: gb_hhg.*

3: gb_in.*

4: gb_om.*

5: gb_ov.*

6: gb_pat.*

7: gb_ph.*

8: gb_pl.*

9: gb_pr.*

10: gb_ro.*

11: gb_sts.*

12: gb_sy.*

13: gb_un.*

14: gb_vi.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2370	100.0	2370	6	AR269117
2	2370	100.0	2370	6	AR487609
3	1665	70.3	2240	6	AX394707
4	1656.2	69.9	1965	6	AX394705
5	1654	69.8	2052	6	AX394703
6	1619	68.3	2249	9	BC063885
7	1449.4	61.2	1947	6	AX166554
8	1349	56.9	1918	6	AX659639
9	1338	56.5	1896	6	AX659636
10	696.2	29.4	3835	6	CQ843384
11	696.2	29.4	3835	9	AX126330
12	602	25.4	1632	6	AX803389
13	597.4	25.2	2281	10	BC053516
14	576.2	24.3	1884	6	CQ732078
15	369.4	15.6	1745	9	BC068201
16	369.4	15.6	5426	6	AX224518
17	369.4	15.6	5448	6	AR338755
18	369.4	15.6	5497	9	AB067488
19	369.4	15.6	5508	6	AX658014

20	369.4	15.6	5511	9	HSM800531
21	369.4	15.6	5532	6	AR338754
22	368	15.5	3645	6	AX224471
23	368	15.5	3729	6	AX658011
24	355.8	15.0	4263	6	AR223879
25	355.8	15.0	4263	10	S45828
26	355.8	15.0	5479	10	AK173292
27	288.6	12.2	958	5	CR386196
28	272.8	11.5	187691	9	AL139082
29	255.2	10.8	2369	5	BC075119
30	234.6	10.7	1578	6	AX657957
31	254.2	10.7	1581	6	AX657962
32	254.2	10.7	2110	6	AX504240
33	254.2	10.7	2134	9	BC019916
34	254.2	10.7	3255	9	AK131359
35	254	10.7	2194	5	BC072363
36	253.8	10.7	876	6	AX688027
37	253.8	10.7	876	6	BD148089
38	253.8	10.7	2975	6	AX658016
39	253.8	10.7	2975	6	AX882436
40	253.8	10.7	2975	6	BD159723
41	253.8	10.7	2975	9	AK027580
42	253	10.7	1997	10	AF093416
43	253	10.7	2025	10	AF099066
44	252.6	10.7	1610	9	AB072828
45	252.6	10.7	2364	6	AX405633

ALIGNMENTS

RESULT 1
AR269117
LOCUS AR269117 2370 bp DNA linear PAT 10-APR-2003
DEFINITION Sequence 1 from patent US 6500656.
ACCESSION AR269117
VERSION AR269117.1 GI:29699967
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 2370)
Webster M., Yan, C., Di Francesco, V. and Beasley, E.
AUTHORS Isolated human kinase proteins, nucleic acid molecules encoding
TITLE human kinase proteins, and uses thereof
JOURNAL Patent: US 6500656-A 1 31-DEC-2002;
FEATURES Location/Qualifiers
source 1..2370
/organism="unknown"
/mol_type="genomic DNA"

ORIGIN

Query Match 100.0%; Score 2370; DB 6; Length 2370;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2370; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 CATTGGAGACCATGGATAGTCAGTGTGATTAAAGGCCATCGGGCAAGGTGCCTTCGGGA 60
Db 1 CATTGGAGACCATGGATAGTCAGTGTGATTAAAGGCCATCGGGCAAGGTGCCTTCGGGA 60
Qy 61 AAGCATACTTAGCTTAAAGGGGAAATCAGATAGCAAGCACTGTGTCTATAAAGAGATCAATT 120
Db 61 AAGCATACTTAGCTTAAAGGGGAAATCAGATAGCAAGCACTGTGTCTATAAAGAGATCAATT 120
Qy 121 TTGAAAGATGCCCATACAGAAAAGAGCTTCAAGAAAGAGTGTCTCTCGAAA 180
Db 121 TTGAAAGATGCCCATACAGAAAAGAGCTTCAAGAAAGAGTGTCTCTCGAAA 180
Qy 181 AGATGAAACATCCCAACATTTAGCTTCTTCAATTCATTTCAGAGAGATGGCAGCTGT 240
Db 181 AGATGAAACATCCCAACATTTAGCTTCTTCAATTCATTTCAGAGAGATGGCAGCTGT 240
Qy 241 TTATTGTATGGAAATATTGTGTGAGGGGATCTCATGAAAAGAGATCATAGACAACGGG 300

LOCUS	AR487609	2370 bp	DNA	linear	PAT 14-MAY-2004
DEFINITION	Sequence 1 from patent US 6706510.				
ACCESSION	AR487609				
VERSION	AR487609.1 GI:47252885				
KEYWORDS					
SOURCE	Unknown.				
ORGANISM	Unknown.				
REFERENCE	Unclassified.				
AUTHORS	1 (bases 1 to 2370)				
TITLE	Webster, M., Yan, C., Di Francesco, V. and Beasley, E.				
JOURNAL	Isolated human kinase proteins				
FEATURES	Patent: US 6706510-A 1 16-MAR-2004;				
source	Location/Qualifiers				
	1..2370				
	/organism="unknown"				
ORIGIN	/mol_type="genomic DNA"				
	Query Match 100.0%; Score 2370; DB 6; Length 2370;				
	Best Local Similarity 100.0%; Pred. No. 0;				
	Matches 2370; Conservative 0; Mismatches 0; Indels 0; Gaps 0;				
Qy	1	CATTGGAGACCATGGATAGTACGATGGATTAAAGGCCATCGGGCAAGTGCTTCGGGA	60		
Db	1	CATTGGAGACCATGGATAGTACGATGGATTAAAGGCCATCGGGCAAGTGCTTCGGGA	60		
Qy	61	AAGCATCTTAGCTAAAGGGAATCAGATAGCAAGCACTGTCTCATAAAGAGATCAATT	120		
Db	61	AAGCATCTTAGCTAAAGGGAATCAGATAGCAAGCACTGTCTCATAAAGAGATCAATT	120		
Qy	121	TTGAAAAGATGCCATACAGAAAAGAGCTTCAAGAAAAGAGTGAATCTTCTGGAAA	180		
Db	121	TTGAAAAGATGCCATACAGAAAAGAGCTTCAAGAAAAGAGTGAATCTTCTGGAAA	180		
Qy	181	AGATGAACATCCCAACATGTAGCTTCTTCATTTCAATTTCAAGAGATGGCAGGCTGT	240		
Db	181	AGATGAACATCCCAACATGTAGCTTCTTCATTTCAATTTCAAGAGATGGCAGGCTGT	240		
Qy	241	TTATTGTAAATGGAATATTGTGATGGAGGGGATCTCATGAAAAGGATCAATAGACACGGG	300		
Db	241	TTATTGTAAATGGAATATTGTGATGGAGGGGATCTCATGAAAAGGATCAATAGACACGGG	300		
Qy	301	GTGTGTTATTAGTGAAGATCAGATCCTCGTTGGTTGTACAGATTTCTCTAGACTAA	360		
Db	301	GTGTGTTATTAGTGAAGATCAGATCCTCGTTGGTTGTACAGATTTCTCTAGACTAA	360		
Qy	361	AACATATTCTATGACAGGAAGATATTACACAGGACATATAAGCTCAGAACATTTTCTTA	420		
Db	361	AACATATTCTATGACAGGAAGATATTACACAGGACATATAAGCTCAGAACATTTTCTTA	420		
Qy	421	GCAAGAACGGAATGTGGCAAGCTTTGGGACTTTGGTATAGCAAGATCCTGAATAATT	480		
Db	421	GCAAGAACGGAATGTGGCAAGCTTTGGGACTTTGGTATAGCAAGATCCTGAATAATT	480		
Qy	481	CCATGGAATCTGCTCGAACTTGATTGGAACACCTTACTCTGCTCCCGAGATCTCTC	540		
Db	481	CCATGGAATCTGCTCGAACTTGATTGGAACACCTTACTCTGCTCCCGAGATCTCTC	540		
Qy	541	AGAATAAACCCCTACAAATAAAGCGATATTGGTCTCTTTGGCTGTCTTATATGAGC	600		
Db	541	AGAATAAACCCCTACAAATAAAGCGATATTGGTCTCTTTGGCTGTCTTATATGAGC	600		
Qy	601	TCTGCACACTTAAACATCTTTTGGGTTACAACTTACAGCAGTGGTCTGAAGATTT	660		
Db	601	TCTGCACACTTAAACATCTTTTGGGTTACAACTTACAGCAGTGGTCTGAAGATTT	660		
Qy	661	GTCAAGCACATTTTGGCCCAATATCTCCGGGGTTTCTCGTGAGCTCCATTCCTTGATAT	720		
Db	661	GTCAAGCACATTTTGGCCCAATATCTCCGGGGTTTCTCGTGAGCTCCATTCCTTGATAT	720		
Qy	721	CTCAGCTCTTTCAAGTATCTCTCGAGACCGACCAATCCATAAATTCATTTGAAAGGC	780		
Db	721	CTCAGCTCTTTCAAGTATCTCTCGAGACCGACCAATCCATAAATTCATTTGAAAGGC	780		

Qy	781	CTTTTGTAGAGATCTTTATCCCAATATTGTGACTCTGAGGTCATTCAGGAAGATTCA	840
Db	781	CTTTTGTAGAGATCTTTATCCCAATATTGTGACTCTGAGGTCATTCAGGAAGATTCA	840
Qy	841	GTCAATGCTTTATATGCAGAGCAGGAGCGCAGCTTCTGCACATGCTGGGAAGGTGCTCC	900
Db	841	GTCAATGCTTTATATGCAGAGCAGGAGCGCAGCTTCTGCACATGCTGGGAAGGTGCTCC	900
Qy	901	AGAAAGTGTAAATAACAAAAGTGGATTTCGGGAAAGTCCCAACCAAGATCAAGGATAT	960
Db	901	AGAAAGTGTAAATAACAAAAGTGGATTTCGGGAAAGTCCCAACCAAGATCAAGGATAT	960
Qy	961	CTGTGCCAATTAAAGGAATGCTATATTGCATAGAAATGAATGGAGACCAACAGCTGGAG	1020
Db	961	CTGTGCCAATTAAAGGAATGCTATATTGCATAGAAATGAATGGAGACCAACAGCTGGAG	1020
Qy	1021	CCAGAAGGCCAGATCTATAAAAATGATAGAAAGACCCAAAATTCGTGCTGTGGAC	1080
Db	1021	CCAGAAGGCCAGATCTATAAAAATGATAGAAAGACCCAAAATTCGTGCTGTGGAC	1080
Qy	1081	ATTATGATTTATTTATGCTCAACTTGTATATGCTGAGGAGGAGGCCCAAAACCAAGTT	1140
Db	1081	ATTATGATTTATTTATGCTCAACTTGTATATGCTGAGGAGGAGGCCCAAAACCAAGTT	1140
Qy	1141	ATCACCCCTATTCTCAAGAAAATACCTGGAGTTTACGGTCAAGGAAACGAGGCATG	1200
Db	1141	ATCACCCCTATTCTCAAGAAAATACCTGGAGTTTACGGTCAAGGAAACGAGGCATG	1200
Qy	1201	GTCCATCCCAAGTCAATGCGCTGTGAGTACCTTCAGAGAAAATTTGAAGCTCAACAAT	1260
Db	1201	GTCCATCCCAAGTCAATGCGCTGTGAGTACCTTCAGAGAAAATTTGAAGCTCAACAAT	1260
Qy	1261	ATAAGTTGAAAGTGGAGAGCAATTCGGTCTTCCTCATCTTCGCGAGCCAAATTACA	1320
Db	1261	ATAAGTTGAAAGTGGAGAGCAATTCGGTCTTCCTCATCTTCGCGAGCCAAATTACA	1320
Qy	1321	ACCAGAGACAAGAGCTTAAGAAGTAATGGAGAGAGCTAGATTTCCAGGAGCTGCCATTTA	1380
Db	1321	ACCAGAGACAAGAGCTTAAGAAGTAATGGAGAGAGCTAGATTTCCAGGAGCTGCCATTTA	1380
Qy	1381	GGAAAACGAAATGAAGGAACAGGAATNTTGGAGCAGTTAGAGGAAATACGCCAACAGT	1440
Db	1381	GGAAAACGAAATGAAGGAACAGGAATNTTGGAGCAGTTAGAGGAAATACGCCAACAGT	1440
Qy	1441	ACCTCAATGACATGAAAGAAATTAGAAAAGATGGGAGAGAACCCAGAGGACATTTGAAA	1500
Db	1441	ACCTCAATGACATGAAAGAAATTAGAAAAGATGGGAGAGAACCCAGAGGACATTTGAAA	1500
Qy	1501	AGAGCTTGAACAAAATGAGGCTTCAGAACACAAAGGAAAGTAAATCCAGAACAGAAAT	1560
Db	1501	AGAGCTTGAACAAAATGAGGCTTCAGAACACAAAGGAAAGTAAATCCAGAACAGAAAT	1560
Qy	1561	ATAAGCTAAGAGGGGGTAAATTTGAAATTAATTTGACAAATGTTATTTCTGATGAA	1620
Db	1561	ATAAGCTAAGAGGGGGTAAATTTGAAATTAATTTGACAAATGTTATTTCTGATGAA	1620
Qy	1621	ACATCTCCCAAGAGGAGGCAATGGATATACCAATGAAACTTTGACCTTTGAGGATG	1680
Db	1621	ACATCTCCCAAGAGGAGGCAATGGATATACCAATGAAACTTTGACCTTTGAGGATG	1680
Qy	1681	GCATGAATTTAAGGAATATGATGTGTAAGAGGAGCATGAGATATACAGACAAGCAT	1740
Db	1681	GCATGAATTTAAGGAATATGATGTGTAAGAGGAGCATGAGATATACAGACAAGCAT	1740
Qy	1741	TTGAAAACCTTCACTCCCAAGAGCAGCATTTACAGAACTGACTTGGCTCAGTTTCTCT	1800
Db	1741	TTGAAAACCTTCACTCCCAAGAGCAGCATTTACAGAACTGACTTGGCTCAGTTTCTCT	1800
Qy	1801	TCCTGGAATATCTCTGCTCATTTTCTCTCTGAAAAAATCTCCATTCAGAGGCATCTTA	1860
Db	1801	TCCTGGAATATCTCTGCTCATTTTCTCTCTGAAAAAATCTCCATTCAGAGGCATCTTA	1860

```
QY 1861 TTGAGGATCTCTTTGTGCGCAACGACTGCTCACTGAGGACTGGAGTGAAGGAATCG 1920
Db 1861 TTGAGGATCTCTTTGTGCGCAACGACTGCTCACTGAGGACTGGAGTGAAGGAATCG 1920
QY 1921 AGCTTAGGACATAACCCCTACCACTACATAAACAACCTTTGGAGAATCAGGAGAGATAA 1980
Db 1921 AGCTTAGGACATAACCCCTACCACTACATAAACAACCTTTGGAGAATCAGGAGAGATAA 1980
QY 1981 GCCAAGGAGAGAGACAGTCACTGAGGAGGACACAGGAATGGCAGATCAACTGGAAGA 2040
Db 1981 GCCAAGGAGAGAGACAGTCACTGAGGAGGACACAGGAATGGCAGATCAACTGGAAGA 2040
QY 2041 GAAAGGCCAGATGAGGGTTTTCCACGCGAGACTGTAGTGTGTTGGGAAACAGGAGGAGT 2100
Db 2041 GAAAGGCCAGATGAGGGTTTTCCACGCGAGACTGTAGTGTGTTGGGAAACAGGAGGAGT 2100
QY 2101 GGGATGAGGAGGCGCTCAGACTCTGTCAGATGATGGCGAGTGGCGACATCACCTCCA 2160
Db 2101 GGGATGAGGAGGCGCTCAGACTCTGTCAGATGATGGCGAGTGGCGACATCACCTCCA 2160
QY 2161 CCTGCCCCACGGGCGCTGACAGTGAAGTCTGCTTGTAGTGTGCTGTCAGGAGGGAAGA 2220
Db 2161 CCTGCCCCACGGGCGCTGACAGTGAAGTCTGCTTGTAGTGTGCTGTCAGGAGGGAAGA 2220
QY 2221 CCAAGGACCCGTACAGCCAGTGTCTATCTGTGATGATGATCTATCTTCTCACTATAC 2280
Db 2221 CCAAGGACCCGTACAGCCAGTGTCTATCTGTGATGATGATCTATCTTCTCACTATAC 2280
QY 2281 CCTATAGATCTGTATCAGACACTTTCAATATGTTGTTTGTATATCTCAGAAATAA 2340
Db 2281 CCTATAGATCTGTATCAGACACTTTCAATATGTTGTTTGTATATCTCAGAAATAA 2340
QY 2341 AAAAAAACAACACTGTATGCGCGTTACGAGCG 2370
Db 2341 AAAAAAACAACACTGTATGCGCGTTACGAGCG 2370

RESULT 3
AX394707 2240 bp DNA linear PAT 18-MAY-2002
DEFINITION Sequence 5 from Patent WO0218555.
ACCESSION AX394707
VERSION AX394707.1 GI:21065808
KEYWORDS Homo sapiens (human)
SOURCE Homo sapiens
ORGANISM Homo sapiens
REFERENCE 1
AUTHORS Friddle,C.J., Hilbun,E., Nepomnichy,B. and Hu,Y.
TITLE Human kinase proteins and polynucleotides encoding the same
JOURNAL Patent: WO 021855-A 5 07-MAR-2002;
Lexicon Genetics Incorporated (US)
FEATURES
source 1..2240
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"

ORIGIN
Query Match 70.3%; Score 1665; DB 6; Length 2240;
Best Local Similarity 94.5%; Pred. No. 0;
Matches 1774; Conservative 0; Mismatches 10; Indels 93; Gaps 1;

QY 1 CATTGGAGACCATGGATAAGTACGATGTGATTAAGGCCATCGGCGAAGTGCCTTCGGGA 60
Db 94 CATTGGAGACCATGGATAAGTACGATGTGATTAAGGCCATCGGCGAAGTGCCTTCGGGA 153
QY 61 AAGCATACTTAGCTAAAGGGGAATCAGATAGCAGCACTGTGTATATAAAGAGATCAATT 120
Db 154 AAGCATACTTAGCTAAAGGGGAATCAGATAGCAGCACTGTGTATATAAAGAGATCAATT 213
QY 121 TTGAAAGATGCCCATACAGAAAGAAAGAGCTTCAAGAAAGAAAGTGAATTTCTCGAAA 180
```

```
Db 214 TTGAAAGATGCCCATACAGAAAGAAAGAGCTTCAAGAAAGAAAGTGAATTTCTTCGAAA 273
QY 181 AGATGAACAATCCCAACATTTGTAGCCTTTCTCAATTCATTTCAAGAGAATGGCAGGCTGT 240
Db 274 AGATGAACAATCCCAACATTTGTAGCCTTTCTCAATTCATTTCAAGAGAATGGCAGGCTGT 333
QY 241 TTATTTAATGGAATATTTGTATGAGGAGGATCTCATGAAAGGATCAATAGACAACGGG 300
Db 334 TTATTTAATGGAATATTTGTATGAGGAGGATCTCATGAAAGGATCAATAGACAACGGG 393
QY 301 GTGTGTTATTTAGTGAAGATCAGATCCTCGGTTGGTTTGTACAGATTTCTCTAGGATAA 360
Db 394 GTGTGTTATTTAGTGAAGATCAGATCCTCGGTTGGTTTGTACAGATTTCTCTAGGATAA 453
QY 361 AACATATTCATGACAGGAAGATATTACACAGGGACATAAAGCTCAGAACATTTTCTTA 420
Db 454 AACATATTCATGACAGGAAGATATTACACAGGGACATAAAGCTCAGAACATTTTCTTA 513
QY 421 GCAAGAACGGGAATGGTGGCAAGCTTTGGGGAATTTGGTATATAGCAAGAGTCTGTAATATT 480
Db 514 GCAAGAACGGGAATGGTGGCAAGCTTTGGGGAATTTGGTATATAGCAAGAGTCTGTAATATT 573
QY 481 CCATGGAACTTTGCTCGAACTTTGTAATGGAAACACTTACTACTCTGCTTGGCTGTCTTATATGAGC 540
Db 574 CCATGGAACTTTGCTCGAACTTTGTAATGGAAACACTTACTACTCTGCTTGGCTGTCTTATATGAGC 633
QY 541 AGAATAAACCCCTCAACAATAAAGCGGATATTTGGTCTCTTGGCTGTCTTATATGAGC 600
Db 634 AGAATAAACCCCTCAACAATAAAGCGGATATTTGGTCTCTTGGCTGTCTTATATGAGC 693
QY 601 TCTGCACACTTTAAACATCTTTTGGGGTAAACAATTTACAGCAGCTGGTTCTGAAGATT 660
Db 694 TCTGCACACTTTAAACATCTTTTGGGGTAAACAATTTACAGCAGCTGGTTCTGAAGATT 753
QY 661 GTCAGACACTTTTGGCCCAATATCTCGGGTTTTTCTCGTGAGCTCCATTTCTTGTATAT 720
Db 754 GTCAGACACTTTTGGCCCAATATCTCGGGTTTTTCTCGTGAGCTCCATTTCTTGTATAT 813
QY 721 CTCAGCTCTTTCAAGTATCTCTCGAGACCGACCATCCATAAATTCATTTTGAAGAGC 780
Db 814 CTCAGCTCTTTCAAGTATCTCTCGAGACCGACCATCCATAAATTCATTTTGAAGAGC 873
QY 781 CCTTTTGAAGAACTTTATTCGAAATATTTGACTCCTGAGGTCATTTCAAGAGAAATTC 840
Db 874 CCTTTTGAAGAACTTTATTCGAAATATTTGACTCCTGAGGTCATTTCAAGAGAAATTC 933
QY 841 GTCACATCTTTATATGACAGCAGGAGCGCCAGCTTTCTCGACATGCTGGGAAGTGTCTCC 900
Db 934 GTCACATCTTTATATGACAGCAGGAGCGCCAGCTTTCTCGACATGCTGGGAAGTGTCTCC 993
QY 901 AGAAGTGTAAATACAAAAGTGAATTCGGGGAAGTGCACCAAGATCAAGAGATAT 960
Db 994 AGAAGTGTAAATACAAAAGTGAATTCGGGGAAGTGCACCAAGATCAAGAGATAT 1053
QY 961 CTGTGCCAATTTAAAGGAATGCTATATTTGCATGAATGMAATGGAGACCAACAGCTGGAG 1020
Db 1054 CTGTGCCAATTTAAAGGAATGCTATATTTGCATGAATGMAATGGAGACCAACAGCTGGAG 1113
QY 1021 CCCAAGAGCCAGATCTATAAATAATGATAGAAAGCCCAAAATTTGCTGTCTGTGGAC 1080
Db 1114 CCCAAGAGCCAGATCTATAAATAATGATAGAAAGCCCAAAATTTGCTGTCTGTGGAC 1173
QY 1081 ATTATGATTTATTTATGCTCAACTGATATGCTGAGGAGAGAGCCCAAAACCAAGTT 1140
Db 1174 ATTATGATTTATTTATGCTCAACTGATATGCTGAGGAGAGAGCCCAAAACCAAGTT 1233
QY 1141 ATCACCTCTATTCTCAAGAAATACTGAGTTGAGGATTAAGGTCAGAGAAACGAGGATG 1200
Db 1234 ATCACCTCTATTCTCAAGAAATACTGAGTTGAGGATTAAGGTCAGAGAAACGAGGATG 1293
QY 1201 GTCCATCCCAAGTCAATGGCTGCTGAGTACCTTCAGAGAAATTTGAGCTCAACAAT 1260
```

```
Db 1294 GTCCATCCCCTCAAGTCAATGGCTGCTGAGTACCTTTCAGAGAAAATTTGAAGCTCAACAT 1353
Qy 1261 ATAAGTTGAAGTGAGAGCAATGGGTCCTTCGTCCTTCCTGCGAGCCAAATTACA 1320
Db 1354 ATAGTTGAAGTGAGAGCAATGGGTCCTTCGTCCTTCCTGCGAGCCAAATTACA 1413
Qy 1321 ACCAGAGACAAGAGCTAAGAGTAATGGAGAGAGCTAGATTCAGAGAGCTGCCATTTA 1380
Db 1414 ACCAGAGACAAGAGCTAAGAGTAATGGAGAGAGCTAGATTCAGAGAGCTGCCATTTA 1473
Qy 1381 GGAAGAACCAATGAAGGACAGGAAATTTGGAAGCAGTTAGAGAAATAGCCCAACAGT 1440
Db 1474 GGAAGAACCAATGAAGGACAGGAAATTTGGAAGCAGTTAGAGAAATAGCCCAACAGT 1533
Qy 1441 ACCTCAATGACATGAAGAAATTTAGAAAGAGATGGGAGAGAACAG 1488
Db 1534 ACCACAATGACATGAAGAAATTTAGAAAGAGATGGGAGAGAACAG 1593
Qy 1489 ----- 1488
Db 1594 AAATAAGTCATTAACCTATTTGGTGAAGAGAGTAACTGCTGTCATCAAGATGCAT 1653
Qy 1489 -----AGGACATTGAAAGAGCTTTGAAACAAATGAGGCTTCAGA 1527
Db 1654 CTGAGGAGAGACACCTGTGCAGGACATTGAAAGAGCTTTGAAACAAATGAGGCTTCAGA 1713
Qy 1528 ACACAAAGGAAAGTAAATTCAGACAGAAATATAAGCTTAAGAGGGGTAAATTTG 1587
Db 1714 ACACAAAGGAAAGTAAATTCAGACAGAAATATAAGAGCTTAAGAGGGGTAAATTTG 1773
Qy 1588 AAATTAATTTAGACAAATGATTTCTGTGATGAAACATCTCCAGAGAGAGGCAATGG 1647
Db 1774 AAATTAATTTAGACAAATGATTTCTGTGATGAAACATCTCCAGAGAGAGGCAATGG 1833
Qy 1648 ATATACCAATGAATTTGACCTTTGAGGATGGATGAAGTTTAAGGAATATGAATGTG 1707
Db 1834 ATATACCAATGAATTTGACCTTTGAGGATGGATGAAGTTTAAGGAATATGAATGTG 1893
Qy 1708 TAAAGGAGCATGGAGATTATACAGACAAAGCAATTTGAAACCTTCACCTGCCAGAGCAG 1767
Db 1894 TAAAGGAGCATGGAGATTATACAGACAAAGCAATTTGAAACCTTCACCTGCCAGAGCAG 1953
Qy 1768 CATTTACAGAACTGACT 1784
Db 1954 GGTTCCTCCAGCAGACT 1970
```

```
RESULT 4
AX394705 1965 bp DNA linear PAT 18-MAY-2002
LOCUS Sequence 3 from Patent WO0218555.
DEFINITION AX394705
ACCESSION AX394705
VERSION AX394705.1 GI:21065807
KEYWORDS Homo sapiens (human)
SOURCE Homo sapiens
ORGANISM Homo sapiens
REFERENCE 1
AUTHORS Friedle C.J., Hilbun E., Nepomnichy B. and Hu Y.
TITLE Human kinase proteins and polynucleotides encoding the same
JOURNAL Patent: WO 0218555-A 3 07-MAR-2002;
Lexicon Genetics Incorporated (US)
FEATURES
Location/Qualifiers
source 1..1965
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
```

ORIGIN

Query Match 69.9%; Score 1656.2; DB 6; Length 1965;
Best Local Similarity 96.1%; Pred. No. 0;
Matches 1710; Conservative 0; Mismatches 63; Indels 6; Gaps 1;

```
Qy 12 ATGATAAGTACGATGTGATTAAGGCCCATCGGCAAGGTGCCTTCGGGAAAGCATACTTA 71
Db 1 ATGATAAGTACGATGTGATTAAGGCCCATCGGCAAGGTGCCTTCGGGAAAGCATACTTA 60
Qy 72 GCTAAGGGAAATCAGATAGCAAGCACTGTGTATATAAGAGAGATCAATTTTGAAGATG 131
Db 61 GCTAAGGGAAATCAGATAGCAAGCACTGTGTATATAAGAGAGATCAATTTTGAAGATG 120
Qy 132 CCATACAGAAAAAAGAGCTTCAAAGAAAAAGAGTGATTTCTTGGAAAAAGATGAACAT 191
Db 121 CCATACAGAAAAAAGAGCTTCAAAGAAAAAGAGTGATTTCTTGGAAAAAGATGAACAT 180
Qy 192 CCCAAATTTAGCTTCTTCAATTTCAAGAGAAATGCGAGGCTGTTTATTTGTAATG 251
Db 181 CCCAAATTTAGCTTCTTCAATTTCAAGAGAAATGCGAGGCTGTTTATTTGTAATG 240
Qy 252 GAATATTTGATGAGGGGATCTCATGAAAAAGATCAATAGACAAACGGGGTGTGTTATTT 311
Db 241 GAATATTTGATGAGGGGATCTCATGAAAAAGATCAATAGACAAACGGGGTGTGTTATTT 300
Qy 312 AGTAAGATCAGATCTCGTTCGTTTGTACAGATTTCTTAGGACTTAAACATATTTAT 371
Db 301 AGTAAGATCAGATCTCGTTCGTTTGTACAGATTTCTTAGGACTTAAACATATTTAT 360
Qy 372 GACAGGAGATATTTACACAGGACATATAAGCTCAGAACATTTTCTTAGCAAGAGCGGA 431
Db 361 GACAGGAGATATTTACACAGGACATATAAGCTCAGAACATTTTCTTAGCAAGAGCGGA 420
Qy 432 ATGTGTCAAAGCTTGGGGACTTTGGTATAGCAAGAGTCTCTGAATTAATTCATGGAAC 491
Db 421 ATGTGTCAAAGCTTGGGGACTTTGGTATAGCAAGAGTCTCTGAATTAATTCATGGAAC 480
Qy 492 GCTCGAATCTGTATTTGAAACACCTTACTACTCTGCTCCAGAGATCTGTGAGATTAACCC 551
Db 481 GCTCGAATCTGTATTTGAAACACCTTACTACTCTGCTCCAGAGATCTGTGAGATTAACCC 540
Qy 552 TACAAACATATAAAGCGGATATTTGGTCTCTTGGCTGTGCTTATATAGAGCTCTGCACACT 611
Db 541 TACAAACATATAAAGCGGATATTTGGTCTCTTGGCTGTGCTTATATAGAGCTCTGCACACT 600
Qy 612 AAACATCTTTGAGGGTAAACAACTTACAGCAGCTGTTCTGAAGATTTGTCAGACACAT 671
Db 601 AAACATCTTTGAGGGTAAACAACTTACAGCAGCTGTTCTGAAGATTTGTCAGACACAT 660
Qy 672 TTTGCCCCAATATCTCGGGGTTTCTCGTGAAGTCCATTCCTTGATATCTCAGCTCTTT 731
Db 661 TTTGCCCCAATATCTCGGGGTTTCTCGTGAAGTCCATTCCTTGATATCTCAGCTCTTT 720
Qy 732 CAAGTATCTCTCGAGACCCACCATCCATAAATTTCCATTTTGAAGGCCCTTTTAGAG 791
Db 721 CAAGTATCTCTCGAGACCCACCATCCATAAATTTCCATTTTGAAGGCCCTTTTAGAG 780
Qy 792 AATCTTATTTCCAAATATTTGACTCTGAGGTCAATTCAGGAAAGATTCAGTCAATGCTT 851
Db 781 AATCTTATTTCCAAATATTTGACTCTGAGGTCAATTCAGGAAAGATTCAGTCAATGCTT 840
Qy 852 ATATGAGAGAGAGGAGCCAGCTTCTCGACATGCTGGGAGGCTGTCAGAGAGTGTAAA 911
Db 841 ATATGAGAGAGAGGAGCCAGCTTCTCGACATGCTGGGAGGCTGTCAGAGAGTGTAAA 900
Qy 912 ATACAAAGATGAGATTTCCGGGAAAGTGCACCAAGATCAAGGATATCTGTGCCAAT 971
Db 901 ATACAAAGATGAGATTTCCAGGGAAAGTGCACCAAGATCAAGGATATCTGTGCCAAT 960
Qy 972 AAAAGGAATGCTATATTTGATAGAAATGAATGAGAGACACACCTGGAGGCCAGAGGCC 1031
Db 961 AAAAGGAATGCTATATTTGATAGAAATGAATGAGAGACACACCTGGAGGCCAGAGGCC 1020
Qy 1032 AGATCTATAAAATGATAGAAAGACCCAAATTTGCTGCTCTGTGTGACATTTATGATAT 1091
Db 1021 AGATCTATAAAATGATAGAAAGACCCAAATTTGCTGCTCTGTGTGACATTTATGATAT 1080
```



```
Db 1021 AGATCTATATAAATGATAGAAAGACCCAAAATTTGCTGTCTGTGGACATTTATGATTAT 1080
Qy 1092 TATTATGCTCAACTTGATATGCTGAGGAGGAGGCCCAACCAAGTTATCACCCCTATT 1151
Db 1081 TATTATGCTCAACTTGATATGCTGAGGAGGAGGCCCAACCAAGTTATCACCCCTATT 1140
Qy 1152 CTTCAAGAAAATACCTGGAGTTGAGGATACGGTTCAGGAAAACGAGGCATGCTCCATCCCA 1211
Db 1141 CCTCAAGAAAATACCTGGAGTTGAGGATACGGTTCAGGAAAACGAGGCATGCTCCATCCCA 1200
Qy 1212 AGTCAATGCGCTGCTGAGTACCTTCAGAGAAAATTTGAAGCTCAACAATAAGTTGAAA 1271
Db 1201 AGTCAATGCGCTGCTGAGTACCTTCAGAGAAAATTTGAAGCTCAACAATAAGTTGAAA 1260
Qy 1272 GTGGAGAACAATTTGGTCTCTGCTCATCTCTGCGAGGCCAAATTAACAACAGAGACAA 1331
Db 1261 GTGGAGAACAATTTGGTCTCTGCTCATCTCTGCGAGGCCAAATTAACAACAGAGACAA 1320
Qy 1332 GAGCTAAGAAGTAATGGAGAGAGCCTAGATTCAGGAGCTGCCATTTAGGAAAAACGAA 1391
Db 1321 GAGCTAAGAAGTAATGGAGAGAGCCTAGATTCAGGAGCTGCCATTTAGGAAAAACGAA 1380
Qy 1392 ATGAGGACACAGATATTTGGAGCAGTTAGAGGAATATACGCCACAGTACCTCAATGAC 1451
Db 1381 ATGAGGACACAGATATTTGGAGCAGTTAGAGGAATATACGCCACAGTACCTCAATGAC 1440
Qy 1452 ATGAAAGAAATTAGAAGAGAGATGGGAGAGAACCCAG----- 1488
Db 1441 ATGAAAGAAATTAGAAGAGAGATGGGAGAGAACCCAGAGGAGAACTCAAAAATTAAGTCAT 1500
Qy 1489 ----- 1488
Db 1501 AAAACCTATTTGGTGAAGAGAGTAACCTGCTGCTCCATCAAGATGCATCTGAGGAGAA 1560
Qy 1489 -----AGGACATTCAGAAAAGACTTTGAAACAAATGAGGCTTCAGAACACAAAGGAA 1538
Db 1561 GCACCTGTGCAGAGCATTCAGAAAAGCTTGAACAAATGAGGCTTCAGAACACAAAGGAA 1620
Qy 1539 AGTAAATTCAGACAGAGAAATATAAGCTAAGAGAGGGGGTAAAAATTTGAAATTAATTTA 1598
Db 1621 AGTAAATTCAGACAGAGAAATATAAGCTAAGAGAGGGGGTAAAAATTTGAAATTAATTTA 1680
Qy 1599 GACAAATGATTTCTGATGABACATCTCCAGAGGAGGAGGCGCATGGATATACCAAT 1658
Db 1681 GACAAATGATTTCTGATGABACATCTCCAGAGGAGGAGGCGCATGGATATACCAAT 1740
Qy 1659 GAACTTTGACCTTTGAGGATGGCATGAAGTTTAAGGAATATGAATGTGTAAAGGAGCAT 1718
Db 1741 GAACTTTGACCTTTGAGGATGGCATGAAGTTTAAGGAATATGAATGTGTAAAGGAGCAT 1800
Qy 1719 GGAGATTATACAGACAAAGCATTGAAAACTTCACTGCCCGAGAGCAGCATTTACAGAA 1778
Db 1801 GGAGATTATACAGACAAAGCATTGAAAACTTCACTGCCCGAGAGCAGCATTTTCCACG 1860
Qy 1779 CTGACT 1784
Db 1861 CAGACT 1866

RESULT 6
LOCUS BC063885
DEFINITION Homo sapiens similar to Serine/threonine-protein kinase Nek1
              (Nima-related protein kinase 1), mRNA (cDNA clone MGC:75495
              IMAGE:30383658), complete cds.
ACCESSION BC063885
VERSION BC063885.1 GI:39645656
KEYWORDS MGC.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
          Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
          Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 2949)
```

AUTHORS

Strausberg, R.L., Feingold, E.A., Grouse, L.H., Derge, J.G., Klausner, R.D., Collins, F.S., Wagner, L., Shenmen, C.M., Schuler, G.D., Altschul, S.F., Zdobych, B., Buettow, K.H., Schaefer, C.F., Bhat, N.K., Hopkins, R.F., Jordan, H., Moore, T., Max, S.I., Wang, J., Hsieh, F., Diatchenko, L., Marusina, K., Farmer, A.A., Rubin, G.M., Hong, L., Stapleton, M., Soares, M.B., Bonaldo, M.F., Casavant, T.L., Schetz, T.E., Brownstein, M.J., Usdin, T.B., Toshiyuki, S., Carninci, P., Prange, C., Raha, S.S., Loquellano, N.A., Peters, G.J., Abramson, R.D., Mullahy, S.J., Bosak, S.A., McEwan, P.J., McKernan, K.E., Malek, J.A., Gunaratne, P.H., Richards, S., Worley, K.C., Hale, S., Garcia, A.M., Gay, L.J., Hulyk, S.W., Vallal, D.K., Muzny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A., Fahey, J., Helton, E., Kettelman, M., Madan, A., Rodriguez, S., Sanchez, A., Whiting, M., Madan, A., Young, A.C., Shevchenko, Y., Bouffard, G.G., Blakeley, R.W., Touchman, J.W., Green, E.D., Dickson, M.C., Rodriguez, A.C., Grimwood, J., Schmutz, J., Myers, R.M., Butterfield, Y.S., Krzywinski, M.I., Skalska, U., Smal, U., Schnerch, A., Schein, J.E., Jones, S.J., and Marra, M.A.

Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences

Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)

12477932

2 (bases 1 to 2949)

Strausberg, R.

Direct Submission

Submitted (08-DEC-2003) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA

NIH-MGC Project URL: <http://mgc.nci.nih.gov>

Contact: MGC help desk

Email: cgapbs-x@mail.nih.gov

Tissue Procurement: Dr. Michael Brownstein

cDNA Library Preparation: Invitrogen Corp

cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

cDNA Sequencing by: Genome Sequence Centre, BC Cancer Agency, Vancouver, BC, Canada

info@bcgsc.bc.ca

Steven Jones, Jennifer Asano, Ian Bosdet, Yaron Butterfield, Susanna Chan, Readman Chiu, Chris Fjell, Erin Garland, Ran Guin, Letticia Hsiao, Martin Krzywinski, Reta Kutche, Oliver Lee, Soo Sen Lee, Victor Ling, Carrie Mathewson, Candice McLeavy, Steven Ness, Pawan Pandoh, Anna-Liisa Prabh, Parvaneh Saedi, Jacqueline Schein, Duane Smal, Michael Smith, Lorraine Spence, Jeff Scott, Michael Thorne, Miranada Tsai, Natasja van den Bosch, Jill Vardy, George Yang, Scott Zuyderduyn, Marco Marra.

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>

Series: IRAC Plate: 141 Row: m Column: 21

This clone was selected for full length sequencing because it passed the following selection criteria: Genomescan gene prediction, similarity but not identity to protein.

FEATURES

source

1. .2949

Location/Qualifiers

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/clone="MGC:75495 IMAGE:30383658"

/tissue_type="Testis"

/clone_lib="NIH_MGC_180"

/lab_host="DH10B"

/note="Vector: pCMV-SPORT6.1"

1. .2949

/gene="MGC75495"

/db_xref="LocusID:341676"

141. .2267

/gene="MGC75495"

/codon_start=1

/product="MGC75495 protein"

/protein_id="AAH63885.1"

/db_xref="GI:39645657"

/db_xref="LocusID:341676"

gene

CDS

/translation="MDKVDVILKALGOGAFGKAYLAKXKSDSKHCVIKBEINPEKMPIOE
KEASKVILILEKXKHNIIVAFNFSFOENGRLFIIVNEYCDGDLKXRIINRQGVLPSE
DQILGVEQISLGLKHLHDKRIILHRDIKAONIPLSNKGMVAKLDFGIARVLNNNSMEL
ARTCIGTPYLSPEIICQKPNYKNTDIWSLGSVLYELCTLKHFPFENNIQOLVLKICQ
AHTFAPISPSRSLISQLFQVSPDRPSINSILSRPLENLIKPYLTPEVIOBEF
SHMLICRAGAPASHAGKVYQKIKIKRQFGKPPRSRISVPIKKNAILHRNEMWPP
AGAOKARSIMWIEPKIAAVCGHYVYVYOLDMLRRRAHKPSVHPIPOENTGVEDYGO
ETRHGPSQSPWPAEYLQKFEAQYKLVKQKOLGLSPSSAEPNYNQBELRSNGEPR
FOELPRKNEKGEYWKQJUEELNQOHNKMEIKRWKNGEPEENSKIHKYLYLVKS
NLPVHQDASEGEAPVQMEFSCCPGWSAMARSLTATSASQDLEKOLKORLQNTKES
KNPEQKYKAKKGVKEINLDKCISDENILQEEEMDIPNETLTFEDGMKFPKYEVCPE
HGDYTDKAFKHLKHPKPEAGFTQTVAAVGNRRQMDGGAPQTLQWMAVADITSTCPTGP
DSESVLSVSRQEGKTKDYPSPVLILM"
147. 917
misc_feature
/gene="MG075495"
/note="S_TKC; Region: Serine/threonine protein kinases,
catalytic domain. Phosphotransferases of the serine or
threonine-specific kinase subfamily. The enzymatic
activity of these protein kinases is controlled by
phosphorylation of specific residues in the activation
segment of the catalytic domain, sometimes combined with
reversible conformational changes in the C-terminal
autoregulatory tail"
/db_xref="CDD:c000180"

ORIGIN

Query Match 68.3%; Score 1619; DB 9; Length 2949;
Best Local Similarity 83.6%; Pred. No. 0;
Matches 1968; Conservative 0; Mismatches 265; Indels 120; Gaps 7;
Qy 1 CATTGGAGACCATGATAAGTACGATGTGATTAAAGCCATCGGCAAGGTGCCTTCGGGA 60
Db 130 CATTGGAGACCATGATAAGTACGATGTGATTAAAGCCATCGGCAAGGTGCCTTCGGGA 189
Qy 61 AAGCATCTTAGCTTAAAGGGAATCAGATAGCAGCACTGTGTATATAAAGAGATCAATT 120
Db 190 AAGCATCTTAGCTTAAAGGGAATCAGATAGCAGCACTGTGTATATAAAGAGATCAATT 249
Qy 121 TTGAAAAGATGCCATCAAGAAAAGAGCTTCAAGAAAAGAGTGAATCTTCTCGAAA 180
Db 250 TTGAAAAGATGCCATCAAGAAAAGAGCTTCAAGAAAAGAGTGAATCTTCTCGAAA 309
Qy 181 AGATGAACATCCCAACATTTGTAGCTTCTCAATTCATTTCAAGAAATGGCAGCTGT 240
Db 310 AGATGAACATCCCAACATTTGTAGCTTCTCAATTCATTTCAAGAAATGGCAGCTGT 369
Qy 241 TTATTGTAATCGAATATTGTATCGAGGGGATCTCATGAAAAGGATCAATAGACAACGGG 300
Db 370 TTAATTGTAATGGAATATTGTATCGAGGGGATCTCATGAAAAGGATCAATAGACAACGGG 429
Qy 301 GTGTGTTATTTAGTGAAGATCAGATCTCGTTGGTTGTACAGATTTCTTAGGACTAA 360
Db 430 GTGTGTTATTTAGTGAAGATCAGATCTCGTTGGTTGTACAGATTTCTTAGGACTAA 489
Qy 361 AACATATTCATGACAGAAGATATTACACAGGGACATATAAAGCTCAGAACATTTTCTTA 420
Db 490 AACATATTCATGACAGAAGATATTACACAGGGACATATAAAGCTCAGAACATTTTCTTA 549
Qy 421 GCAAGAACGGAAATGGTGGCAAGCTTCGGGACTTTGGTATAGCAAGAGTCTCTGAATAAT 480
Db 550 GCAAGAACGGAAATGGTGGCAAGCTTCGGGACTTTGGTATAGCAAGAGTCTCTGAATAAT 609
Qy 481 CCATGGAACTTGCTCGAACTTGATTGGAACACCTTACTACTCTGTCGCCAGAGATCTGTC 540
Db 610 CCATGGAACTTGCTCGAACTTGATTGGAACACCTTACTACTCTGTCGCCAGAGATCTGTC 669
Qy 541 AGAATAAACCCCTACAAACATATAAAGCGATATTTGGTCTTCTTGGCTGTCTTATATGAGC 600
Db 670 AGAATAAACCCCTACAAACATATAAAGCGATATTTGGTCTTCTTGGCTGTCTTATATGAGC 729
Qy 601 TCTGCACATTTAAACATCTTTTGAGGGTAACTTACAGCACTTACAGCACTGTTCTCAAGATT 660
Db 730 TCTGCACATTTAAACATCTTTTGAGGGTAACTTACAGCACTTACAGCACTGTTCTCAAGATT 789

Qy 661 GTCAAGGACATTTGGCCCCAATATCTCGGGGTTTTCTCGTGAGCTCCATTCCTTGATAT 720
Db 790 GTCAAGGACATTTGGCCCCAATATCTCGGGGTTTTCTCGTGAGCTCCATTCCTTGATAT 849
Qy 721 CTCAGCTCTTTCAAGATATCTCTCGAGACCCACCACCATCAATAAATTCATTTGAAAAGGC 780
Db 850 CTCAGCTCTTTCAAGATATCTCTCGAGACCCACCACCATCAATAAATTCATTTGAAAAGGC 909
Qy 781 CTTTTTATAGAAATCTTTATTCGAAAATTTTGACTCTCTGAGGTCAATTCAGAAAGAAATCA 840
Db 910 CTTTTTATAGAAATCTTTATTCGAAAATTTTGACTCTCTGAGGTCAATTCAGAAAGAAATCA 969
Qy 841 GTCACATGCTTATATGACAGACGAGCGCCAGCTTCTCGACATGCTGGGAAGGTGGTCC 900
Db 970 GTCACATGCTTATATGACAGACGAGCGCCAGCTTCTCGACATGCTGGGAAGGTGGTCC 1029
Qy 901 AGAAGTGTAAATAACAAAAGTGAGATTCCGGGAAAGTGCACCAAGATCAAGGATAT 960
Db 1030 AGAAGTGTAAATAACAAAAGTGAGATTCCAGGGAAAGTGCACCAAGATCAAGGATAT 1089
Qy 961 CTGTGCCAATTAAAGGAATGCTATTTTGCATAGAAATGATGAGACCCACCCTGGAG 1020
Db 1090 CTGTGCCAATTAAAGGAATGCTATTTTGCATAGAAATGATGAGACCCACCCTGGAG 1149
Qy 1021 CCCAGAGCGCAGATCTATAAAAAATGATAGAAAGACCCAAAATTTGCTGCTGTCTGGAC 1080
Db 1150 CCCAGAGCGCAGATCTATAAAAAATGATAGAAAGACCCAAAATTTGCTGCTGTCTGGAC 1209
Qy 1081 ATTATGATTTATTTATGCTCAACTTGTATGCTGAGGAGGAGAGCCACAAAACCAAGTT 1140
Db 1210 ATTATGATTTATTTATGCTCAACTTGTATGCTGAGGAGGAGAGCCACAAAACCAAGTT 1269
Qy 1141 ATCAACCTTATTCCTCAAGAAAATACCTGGAGTTGAGGATTACGGTCAGSAAACGAGGATG 1200
Db 1270 ATCAACCTTATTCCTCAAGAAAATACCTGGAGTTGAGGATTACGGTCAGSAAACGAGGATG 1329
Qy 1201 GTCCATCCCAAGTCAATGGCTGCTGAGTACCTTCAGAGNAATTTGAACTTCAACAAT 1260
Db 1330 GTCCATCCCAAGTCAATGGCTGCTGAGTACCTTCAGAGNAATTTGAACTTCAACAAT 1389
Qy 1261 ATAACTTCAAAAGTGGAGAGCAATTTGGGTCTTCCTCATCTTCTGCCGAGCCAAATATACA 1320
Db 1390 ATAACTTCAAAAGTGGAGAGCAATTTGGGTCTTCCTCATCTTCTGCCGAGCCAAATATACA 1449
Qy 1321 ACCGAGACAAGAGCTAAGAAAGTAATGAGAAAGCCCTAGATTCCAGAGGTGCTCATTTA 1380
Db 1450 ACCGAGACAAGAGCTAAGAAAGTAATGAGAAAGCCCTAGATTCCAGAGGTGCTCATTTA 1509
Qy 1381 GAAAAACGAAATGAAGGAACAGGAATATTGGAAGCAGTTAGAGGAAATACGCCAACAGT 1440
Db 1510 GAAAAACGAAATGAAGGAACAGGAATATTGGAAGCAGTTAGAGGAAATACGCCAACAGT 1569
Qy 1441 ACCTCAATGACATCAAAAGAAATTTAGAAAGAGATGGGAGAGAAACCGAGAGCAATTGAAA 1500
Db 1570 ACCCAATGACATCAAAAGAAATTTAGAAAGAGATGGGAGAGAAACCGAGAGCAATTGAAA 1629
Qy 1501 AAGACTTGAACAAATGAGGCTTCAGAACACAAAGGAAGTAAAAATCCAGAACAGAAAT 1560
Db 1630 AATTAAGTCAATAAACCTTATTGTTGTTG-----AAGAAAGATTAACCTGCTTCCATCAAG 1683
Qy 1561 ATAAAGCTTAAAGAGGGGTAAAAATTTGAAATTTAATTTAGACAAATGTATTCTTGATGAAA 1620
Db 1684 ATGCAATCTGAGGAGAGAACCACTGTGACAGATGGAATTTGCTCTTGTGCCCCAG----- 1737
Qy 1621 ACATCCTCAAGAGAGAGAGCAATGGATATACAAATGAAACTTTGACCTTTGAGGATG 1680
Db 1738 -----GCTGGAGTGCATGGCAGCATCTTGGCTCAACGCAACCTCCGCTTCCAGGA-- 1789
Qy 1681 GCATGAAGTTTAAAGGAATATGAATGTGTAAAGGAGCATGGAGATTATACAGACAAAGCAT 1740
Db 1790 -----CATTGAAAAGAGCTTGAAACAAATGAGGCTTCAGAACACAAAGGAA 1835

Db	1008	TTTCCAGATAAAATGATAGAAAGACCCAAAATTGCTGCTGTGTGGACATTATGATTA	1067
Qy	1091	TTATTATGCTCAACTTGATATGCTGTAGGAGGAGCCCAACAACCAAGTTATCACCCCTAT	1150
Db	1068	TTATTATGCTCAACTTGATATGCTGTAGGAGGAGCCCAACAACCAAGTTATCACCCCTAT	1127
Qy	1151	TCCTCAAGAAAAATACATCGGAGTTACGGTTCAGGAAACGAGGCATGGTCCATCCCC	1210
Db	1128	TCCTCAAGAAAAATACATCGGAGTTACGGTTCAGGAAACGAGGCATGGTCCATCCCC	1187
Qy	1211	AAGTCAATGGCCCTGCTGAGTACCTTCAGAGAAATTTGAAGCTCAACAATATAAGTTGAA	1270
Db	1188	AAGTCAATGGCCCTGCTGAGTACCTTCAGAGAAATTTGAAGCTCAACAATATAAGTTGAA	1247
Qy	1271	AGTGGAGAGACCAATTGGGTTCTTCGTCCATCTTCTGCGGAGCCAAATTTACAACGAGAGACA	1330
Db	1248	AGTGGAGAGCAATT-----	1262
Qy	1331	AGAGCTAAGAAGTAATGGAGAGAGCGCTAGATTCCAGGAGCTGCCATTTAGGAAAACGA	1390
Db	1263	-----	1262
Qy	1391	AATGAAGACACGAATATTTGGAAGCAGTTAGAGGAAATACGCCAACAGTACCTCAATGA	1450
Db	1263	-----GGAATATTGGAAGCAGTTAGAGGAAATACGCCAACAGTACCAATGA	1310
Qy	1451	CATGAAAGAAATTAGAAAAGAGATGGGGAGAGAACGAGAGGACATTGAAAAGACTTGAA	1510
Db	1311	CATGAAAGAAATTAGAAAAGAGATGGGGAGAGAACGAGAGGACATTGAAAAGACTTGAA	1370
Qy	1511	ACAAATGAGCCTTCAGAACACAAAGGAAAGTAAATCCAGACAGAGAAATATAAAGCTAA	1570
Db	1371	ACAAATGAGCCTTCAGAACACAAAGGAAAGTAAATCCAGACAGAGAAATATAAAGCTAA	1430
Qy	1571	GAAGGGGGTAAAAATTTGAAATTAATTTAGACAAATGTATTTCTGATGAAAACATCTCTCA	1630
Db	1431	GAAGGGGGTAAAAATTTGAAATTAATTTAGACAAATGTATTTCTGATGAAAACATCTCTCA	1490
Qy	1631	AGAGGAGAGGCAATGGATATACCAATGAAACTTTGACCTTTGAGGATGGCATGAAGTT	1690
Db	1491	AGAGGAGAGGCAATGGATATACCAATGAAACTTTGACCTTTGAGGATGGCATGAAGTT	1550
Qy	1691	TAAGGAATATGAATGTCTAAAGGAGCATGGAGATTATACAGACAAAGCAATTGAAAAA	1750
Db	1551	TAGGAATATGAATGTCTAAAGGAGCATGGAGATTATACAGACAAAGCAATTGAAAAA	1610
Qy	1751	TCAGTCCCGAGAGCAGCATTTACAGAACTGACT	1784
Db	1611	TCAGTCCCGAGAGCAGGGTTTTCCACGCAGACT	1644
RESULT 10			
LOCUS	CQ843384	3635 bp	DNA
DEFINITION	Sequence 2031 from Patent EP1440981.		linear
ACCESSION	CQ843384		PAT 02-AUG-2004
VERSION	CQ843384.1		
KEYWORDS	GI:50895171		
SOURCE	Homo sapiens (human)		
ORGANISM	Homo sapiens		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
AUTHORS	Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.		
TITLE	Isogai, T., Sugiyama, T., Otsuki, T., Wakamatsu, A., Sato, H., Ishii, S.,		
JOURNAL	Yamamoto, J., Isono, Y., Nagai, K. and Irie, R.		
FEATURES	Full-length human cdna		
source	Patent: EP 1440981-A 2031 28-JUL-2004;		
	Research Association for Biotechnology (JP)		
	Location/Qualifiers		
	1..3635		
	/organism="Homo sapiens"		
	/mol type="unassigned DNA"		

Kodaira,H., Furuya,T., Takahashi,M., Kikkawa,E., Omura,Y., Abe,K.,
Kamihara,K., Katsuta,N., Sato,K., Tanikawa,M., Yamazaki,M.,
Sugiyama,T., Irie,R., Otsuki,T., Sato,H., Ota,T., Wakamatsu,A.,
Ishii,S., Yamamoto,J., Isono,Y., Kawai-Hio,Y., Saito,K.,
Nishikawa,T., Kimura,K., Yamashita,H., Matsuo,K., Nakamura,Y.,
Sekine,M., Kikuchi,H., Murakawa,K., Kanehori,K.,
Takahashi-Fujii,A., Oshima,A., Sugiyama,A., Kawakami,B., Suzuki,Y.,
Sugano,S., Nagahara,K., Masuho,Y., Nagai,K. and Isogai,T.
NEBO human cDNA sequencing project
Unpublished
2 (bases 1 to 3635)
Isogai,T. and Yamamoto,J.
Direct Submission
Submitted (15-JUL-2003) Takao Isogai, FLJ Project (HRI Team); 2-6-7
Kazusa-Kamatari, Kisarazu, Chiba 292-0818, Japan
(E-mail:genomics@hri.co.jp, Tel:81-438-52-3975, Fax:81-438-52-3986)
NEBO human cDNA sequencing project supported by Ministry of
Economy, Trade and Industry of Japan; CDNA full insert sequencing;
Research Association for Biotechnology (RAB); cDNA library
construction; Helix Research Institute (HRI) (supported by Japan
Key Technology Center etc.); 5'- & 3'-end one pass sequencing; RAB,
HRI, and Biotechnology Center, National Institute of Technology and
Evaluation; clone selection for full insert sequencing; HRI and
RAB; annotation: HRI and RAB.

TITLE
JOURNAL
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT

FEATURES
source

1. .3635
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="TRACH300689"
/tissue type="trachea"
/clone_lib="TRACH3"
/note="cloning vector: pME18SFL3"
2172. .3470
/note="unnamed protein product"
/codon_start=1
/protein_id="BAC86527.1"
/db_xref="GI:34532789"
/translation="MSCWHLDPGVAGQPSLPSRGRAEAPLTSRRLAGRAD
PPPSRGRGLAGRADPPPSRGRRLAWRLTSLDGVAAAGRRSSPPRGGCR
ADRLTSGTGRUPGGAGPHFDGAVARRVSSLLRWGAEATLLTSGTSPRGAPH
IPDGAAGRRSPHLRWAAGQRSSLPWMDGGAEATLLTGTGPGRGAPHPDDGR
GRDAPHFDPGVAAGQRLQSRHFGRPRQAGRWRL"

CDS

ORIGIN

Query Match 29.4%; Score 696.2; DB 9; Length 3635;
Best Local Similarity 98.6%; Pred. No. 4.4e-152;
Matches 712; Conservative 0; Mismatches 8; Indels 1; Gaps 1;

782 CTTTATAGAGATCTTATCCCAATATTTGACTCTCTGAGGTCATTGAGGAAGATTTCAG 841
1 CTTTATAGAGATCTTATCCCAATATTTGACTCTCTGAGGTCATTGAGGAAGATTTCAG 60
842 TCACATGCTTATATGACGAGCAGGAGCGCCAGCTTCTGACATGCTGGAAAGGTGGTCCA 901
61 TCACATGCTTATATGACGAGCAGGAGCGCCAGCTTCTGACATGCTGTGAAGGTGGTCCA 120
902 GAAGTGTAAATACAAAAGTCAGATTTCGGGGAAGTCCGACCAAGATCAAGGATATC 961
121 GAAGTGTAAATACAAAAGTCAGATTTCGAGGAAGTCCGACCAAGATCAAGGATATC 180
962 TGTGCAATTAAGAAGATGCTTATATGATGAATGAATGAATGAGACCAACAGCTGGAGC 1021
181 TGTGCAATTAAGAAGATGCTTATATGATGAATGAATGAATGAGACCAACAGCTGGAGC 240
1022 CCAGAGGCGCAGATCTTAAATATGATGAAGACCAATTTGCTGTCTGTGGACA 1081
241 CCAGAGGCGCAGATCTTAAATATGATGAAGACCAATTTGCTGTCTGTGGACA 300
1082 TTATGATTATTATTATGCTCACTTGTATGCTGAGGAGGAGGCCCAACCAAGTTA 1141
301 TTATGATTATTATTATGCTCACTTGTATGCTGAGGAGGAGGCCCAACCAAGTTA 360

1142 TCACCTATTCTCTCAAGAAATATCTGGAGTTGAGGATTACGGTCAGGAAACGAGGCATGG 1201
361 TCACCTATTCTCTCAAGAAATATCTGGAGTTGAGGATTACGGTCAGGAAACGAGGCATGG 420
1202 TCCATCCCAAGTCAATGGCCTGCTGAGTACCTTTTCAGAGAAATTTGAAGCTCAACATA 1261
421 TCCATCCCAAGTCAATGGCCTGCTGAGTACCTTTTCAGAGAAATTTGAAGCTCAACATA 480
1262 TAAGTTGAAAGTGGAGAGCAATTTGGGTCTTCGTCCTCATCTTCTCCGAGCCAAATTTACAA 1321
481 TAAGTTGAAAGTGGAGAGCAATTTGGGTCTTCGTCCTCATCTTCTCCGAGCCAAATTTACAA 540
1322 CCAGAGCAAGAGCTAAGAAAGTAATGAGAGAGAGCCCTAGATTCCAGAGCTGCCATTTAG 1381
541 CC-GAGAGCAAGAGCTAAGAAAGTAATGAGAGAGAGCCCTAGATTCCAGAGCTGCCATTTAG 599
1382 GAAAGCAAGATGAAGAAACAGGAATATTGGAAGCACTTAGAGGAAATACGCCAACAGTA 1441
600 GAAAGCAAGATGAAGAAACAGGAATATTGGAAGCACTTAGAGGAAATACGCCAACAGTA 659
1442 CCTCAATGACATGAAGAAATTTAGAAAGAGATGGGAGAGAACCCAGAGGACATTTGAAA 1501
660 CCACATGACATGAAGAAATTTAGAAAGAGATGGGAGAGAACCCAGAGGAGACTCAA 719
1502 A 1502
720 A 720

RESULT 12

AX803389 AX803389 1632 bp DNA linear PAT 24-NOV-2003
LOCUS Sequence 51 from Patent WO03000901.
DEFINITION AX803389
ACCESSION AX803389
VERSION AX803389.1 GI:38502052
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1
AUTHORS Martinez,R.A. and Sigurdsson,G.T.
TITLE Nucleic acids encoding protein kinases
JOURNAL Patent: WO 03000901-A 51 03-JAN-2003;
Decode Genetics EHP. (IS)

FEATURES
source

1. .1632
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
1. .1632
/note="MOOSE03455"
/codon_start=1
/product="OPK V Nima family"
/protein_id="CAE55983.1"
/db_xref="GI:38502053"
/translation="YRNVSRARHLRSLVGLGVFLRGVAYRAKRRERDEWLLSGAVK
KHALIKNADKILKSFQHMIFLTENGLFVMEYCDGDLMKRINRQGVLFSE
QILGWFVQISLGLKHIHDKILRDIKAQNIFLSKNGMVAKLGDGFGARVLANSMELA
RTCTGTFYVLSPEICQNKFNKNTDIWSGLVLYELCTLKHFFEGNNLQQLVLIQOA
HFAPISGFSRELHLSISQIFQVPSRDRPSINSILKPFLENLIPKLTPEVSPFYLQ
RKPEAQYKLVKELQVSVKYLRLNSAVELILKLLKAPHKFTLPSPKVSRVGT
DHCFPRASHETASLVAQPHMAGEASQSQWKAKEOSHALHGRQESLWELLFTKPSD
LLKAAVSHDCATVPQGIQSKTRSKNALLTSOTGRGKGAPHISDDGRPRGDARHP
LDGMAAGRRSSLSRLGSQAEGLTTSQTMGGAETLLTSQTLGRQENCLNLTGGGCGSE
PRSSGRGHPCLVFPVFKGNASSFCFPKNDKVRWR"

ORIGIN

Query Match 25.4%; Score 602; DB 6; Length 1632;
Best Local Similarity 98.4%; Pred. No. 5.1e-130;
Matches 608; Conservative 0; Mismatches 10; Indels 0; Gaps 0;
215 TTCAATTCAGAGATGGCAGGCTTTTATTGTATGTAATGGAATATTGTGATGGAGGGATCT 274

Db	201																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																												</
----	-----	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	----

Butterfield, V.S., Krzywinski, M.I., Skalska, U., Smailus, D.E.,
Schnerch, A., Schein, J.B., Jones, S.J. and Marra, M.A.
Generation and initial analysis of more than 15,000 full-length
human and mouse cDNA sequences
Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
22389257
12477932
2 (bases 1 to 2281)
Strausberg, R.
Direct Submission
Submitted (09-JUN-2003) National Institutes of Health, Mammalian
Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
USA
NIH-MGC Project URL: <http://mgc.nci.nih.gov>
Contact: MGC help desk
Email: cgapbs@mail.nih.gov
Tissue Procurement: Susan L. Sullivan, PhD.
CDNA Library Preparation: ResGen, Invitrogen Corp
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: National Institutes of Health Intramural
Sequencing Center (NISC),
Gaithersburg, Maryland;
Contact: nisc.mgc@nih.gov
Web site: <http://www.nisc.nih.gov/>
Akteer, N., Avele, K., Beckstrom-Sternberg, S.M., Benjamin, B.,
Blakesley, R.W., Bouffard, G., Breen, K., Brinkley, C., Brooks, S.,
Dierich, N.L., Granite, S., Guan, X., Gupta, J., Hachighi, P.,
Hansen, N., Ho, S.-L., Karlins, E., Kwong, P., Laric, P., Legaspi, R.,
Maduro, Q.L., Masiello, C., Maskeri, B., Mastrian, S.D., McCloskey, J.C.,
McDowell, J., Pearson, R., Stantripop, S., Thomas, P.J., Touchman, J.W.,
Tsurgeon, C., Vogt, J.L., Walker, M.A., Wetherby, K.D., Wiggins, L.,
Young, A., Zhang, L.-H. and Green, E.D.
Clone distribution: MGC clone distribution information can be found
through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
Series: IRAK Plate: 100 Row: d Column: 16
This clone was selected for full length sequencing because it
passed the following selection criteria: Hexamer frequency ORF
analysis, similarity but not identity to protein.
Location/Qualifiers
1..2281
/organism="Mus musculus"
/mol_type="mRNA"
/db_xref="taxon:10090"
/clone="MGC:54821 IMAGE:6311513"
/tissue_type="Olfactory epithelium, neonatal mouse,
C57BL/6"
/clone_lib="NIH_MGC_129"
/lab_host="DH10B"
/note="Vector: pCMV-SPORT6"
1..2281
/gene="BB049667"
/note="synonym: 6530411J22"
/db_xref="LocusID:330721"
/db_xref="MGI:2142824"
78..1961
/gene="BB049667"
/codon_start=1
/product="BB049667 protein"
/protein_id="AAH53516.1"
/db_xref="GI:31565170"
/db_xref="LocusID:330721"
/db_xref="MGI:2142824"
/translation="MDNFHLIKIGESTFGKYLAKOKSESSHCIVKIEISITKKEAS
KNEVTLARVEHPNITVFSSQNGRLFIYVEYCDGLMQRIGRQGVNFSDDQIL
CWFQVLSGLKHIDKRLDKSQNI FLKNGWAKLGDGFTARTLNDMSLQAQTC
AGTPYLSPEI CONRPYNNKTDIWSLGVLYELCTLKHFFSNPNPHILVLCQGRVA
PIPSHSDLOSILPOLFRVSPQPSVTSLLKRPFLTLIARLSSVEPFPNBEKLO
HNENAIPTACWRSVPSAAVLOKPEAQOYKLKVERQLGRPSVVEPFPNBEKLO
SHNEETKFOELQYRNKWDQYWKQLEIRQOYHNDKKEIKKMGRELKRVKFEIS
LDKCSSEDTVOENSAVDKLNATLSFDGTFRQSHRCKHEHEDYTDRAFEELCPGEAE
GFFQDVIAENRQWDAGAPHTLLRIMANADVTSTCTPMPDDGQVIVMEGSEVNGKQW

WLDVPTPCALAAECACSGSLASKEGTWIKPQLPKEDQEKVEIATGIMVDEQLEP
GSDDDIKFESEDELSEIIESLEKLAASTEARQAPSSSKVABEPGEKNTLPVK
KIQ"
misc_feature 87..842
/gene="BB049667"
/note="S_TKc; Region: Serine/Threonine protein kinases,
catalytic domain."
/db_xref="CDD:smart00220"

ORIGIN

Query Match 25.2%; Score 597.4; DB 10; Length 2281;
Best Local Similarity 81.1%; Pred. No. 6.1e-129;
Matches 724; Conservative 0; Mismatches 156; Indels 13; Gaps 2;
QY 10 CCATGATAGTACGATGTGATTAGGCCATCGGCAAGTCCCTTCGGGAAAGCATACT 69
DB 76 CCATGATAGTACCTTACCTGATTAGATCATTTGGGAGGCCACCTTTGGGAGGTGTA 135
QY 70 TAGCTAAAGGGAATCAGATACAGACACTGTGTCTATAAAGAGATCAATTTGAAAGA 129
DB 136 TGCTAAAGATTAATCAGAAAGCATCTGTGTCTATAAAGAAATTCAGTTTGAAGA- 194
QY 130 TGCCCATACAAGAAAGAGCTTCAAAGAAAGAGTATTCTTCTGGAAGAGATGAAC 189
DB 195 -----GAAAGAGAGGCTCAAAGAACAGTGTATTCTTCTGGCTAGATGGAGC 243
QY 190 ATCCCAACATTGTAGCTTCTCAATTCAATTTCAAGAGAAATGCGAGGCTGTTATTGTAA 249
DB 244 ATCCCAATATCGTAACCTTCTTCAGCTCGTTTCAAGAGAACGCGAGGCTGTTATTGTAA 303
QY 250 TGGAAATATTGTATGAGGGGATCTCATGAAAAGGATCAATAGACAAAGGGGTGTATT 309
DB 304 TGGAACTGTGTATGAGGGGATCTCATGACAGAGATCCAGAGGAGCGGGGAGTGT 363
QY 310 TTATGTAGATCAGATCTCGGTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 369
DB 364 TCAGGGAAGACCATCTCTGTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 423
QY 370 ATGACAGGAAGATATTACACAGGACATAAAGCTCAGAACATTTTCTTAGCAGAGAG 429
DB 424 ATGACAGGAAGATTTTACACAGGACATAAATCTCAGAAATTTTCTTAGCAGAGATG 483
QY 430 GAATGGTGAAGCTTGGGACTTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 489
DB 484 GAATGGTGAAGCTTGGGACTTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 543
QY 490 TTGCTCGAATTTGTTGGAACACCTTACTACTGCTCCAGAGATCTGTCAGATTAAC 549
DB 544 TTGCTCAACATGTGCTGGGACCTTACTACTGCTCCAGAGATCTGTCAGAACAGGC 603
QY 550 CCTCAACAAATAAAGCGGATATTTGGTCTCTGGGCTGTCTTATATGAGCTCTGCACAC 609
DB 604 CATACAAATAAAGCGGACATCTGGTCTCTGGGCTGTCTTATATGAGCTCTGCACAC 663
QY 610 TTAACATCTCTTGAAGGTAAACAATTACAGAGCTGGTCTGGAAGATTTGTCAAGCAC 669
DB 664 TCAAGCATCTCTTGAAGGTAAACAATTACAGAGCTGGTCTGGAAGATTTGTCAAGCAC 723
QY 670 ATTTTGGCCCAATATCTCGGGGTTTCTCGTGAGCTCCATTCCTTGATATCTCAGCTCT 729
DB 724 GTTGTCTCCCATATACCCCACTTCTCTGAGACCTACAGTCTTGATACCTCAGCTCT 783
QY 730 TTCAAGTATCTCTCGAGACCGACCATCAATAAATTCATTTTGAAGGCGCTTTTATAG 789
DB 784 TCAGAGTGTCTCTCAGGACCGGACCTCGTTACGTCTGCTTTGAAAAGACCCCTTTTATG 843
QY 790 AGAATCTTATCCCAATATTGATCTCTGAGGTC-ATTCAGGAGAAATTCAGTCAATG 848
DB 844 AAATCTCTATGCGCGATCTTGTATCTCTGAGGTCGTTTCAAGAAAGAAATTCAGTCCACG 903
QY 849 CTTATATGACAGAGCGGCGGACCTTCTCGACATGCTGGGAAGGTGTTGTTGTTGTTGTT 901
DB 904 CTCACATGGAGAACATGGCCATTGGCCCAACAGCTTGTGTTGAGAGTTAGTCCA 956

RESULT 14

CQ732078 1884 bp DNA linear PAT 03-FEB-2004
Sequence 19012 from Patent WO02068579.
CQ732078
CQ732078.1 GI:42310392
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1
Venter, C.J., Adams, M.C., Li, P.W. and Myers, E.W.
Kits, such as nucleic acid arrays, comprising a majority of
humanexons or transcripts, for detecting expression and other uses
thereof: WO 02068579-A 18012 06-SEP-2002;
PE Corporation (NY) (US)
Location/Qualifiers
1. 1884
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
Query Match 24.3%; Score 576.2; DB 6; Length 1884;
Best Local Similarity 59.3%; Pred. No. 5.6e-124;
Matches 1017; Conservative 0; Mismatches 123; Indels 327; Gaps 4;
QY 318 GATCAGATCTCGGTGGTGGTTGTACAGATTTCTCTAGGACTAAACATATTCATGACGG 377
DB 112 GAAAGATCTCGGTGGTGGTTGTACAGATTTCTCTAGGACTAAACATATTCATGACGG 171
QY 378 AAGATATTACACAGGGACATAAAGCTCAGAACATTTTCTTAGCAAGAACGGAATGGTG 437
DB 172 AAGATATTACACAGGGACATAAAGCTCAGAACATTTTCTTAGCAAGAACGGAATGGTG 231
QY 438 GCAAAGCTGGGACCTTTGGTATAGCAAGAGTCTGTAATTAATTCATGGAACCTTGTCTGA 497
DB 232 GCAAAGCTGGGACCTTTGGTATAGCAAGAGTCTGTAATTAATTCATGGAACCTTGTCTGA 272
QY 498 ACTTGTATTGGAACACCTTACTACTCTGCCAGAGATCTGTGAGATAAACCCTACAAC 557
DB 273 ----- 272
QY 558 AATAAAACGATATTGGTCTCTTGGCTGTCTTATATGAGCTCTGCACACTTAAACAT 617
DB 273 -----GGATATTGGTCTCTTGGCTGTCTTATATGAGCTCTGCACACTTAAACAT 324
QY 618 CCTTTTGAGGTAACACTTACAGCAGCTGGTCTGAGATTTGTCAAGCACATTTTGGC 677
DB 325 CCTTTTGAGGTAACACTTACAGCAGCTGGTCTGAGATTTGTCAAGCACATTTTGGC 384
QY 678 CCAATATCTCCGGGTTTCTCGTGGCTCCATTCCTTGATATCTCAGCTCTTTCAGTA 737
DB 385 CCAATATCTCCGGGTTTCTCGTGGCTCCATTCCTTGATATCTCAGCTCTTTCAGTA 444
QY 738 TCTCCTCGAGACCGACATCCATTAATTCATTTTGAAGGCGCTTTTGTAGAAATCTTT 797
DB 445 TCTCCTCGAGACCGACATCCATTAATTCATTTTGAAGGCGCTTTTGTAGAAATCTTT 504
QY 798 ATTCCCAAAATATTGACTCTGAGGTCATTTCAGGAAGATTCAGTCCATGCTTATATGC 857
DB 505 ATTCCCAAAATATTGACTCTGAG----- 528
QY 858 AGAGCAGGAGCGGACGCTTCTCGACATGCTGGGAAGGTGTCAGAAAGTGTAAATAACAA 917
DB 529 ----- 528
QY 918 AAAGTGAGATTCCGGGAAAGTCCCAAGATCAAGGATATCTGTGCCAATTTAAAGG 977

```
Db 529 ----- 528
Qy 978 AATGCTATATTGCATAGAAATGAATGGAGACACCAGCTGGAGCCCGAGAGCCAGATCT 1037
Db 529 ----- 528
Qy 1038 ATAAAAATGATAGAAAGACCCAAATATGCTGCTGTGTGCGACATATGATATATTAT 1097
Db 529 ATAAAAATGATAGAAAGACCCAAATATGCTGCTGTGTGCGACATATGATATATTAT 588
Qy 1098 GCTCAACTGATATGCTGAGGAGGAGGCCACAAACCAAGTTATCACCTATTCCTCAA 1157
Db 589 GCTCAACTGATATGCTGAGGAGGAGGCCACAAACCAAGTTATCACCTATTCCTCAA 648
Qy 1158 GAAATATCTGAGTTGAGGATTACGGTTCAGGAACGAGGCGATGGTCCATCCCAAGTCAA 1217
Db 649 GAAATATCTGAGTTGAGGATTACGGTTCAGGAACGAGGCGATGGTCCATCCCAAGTCAA 708
Qy 1218 TGGCTGCTGAGTACCTTCAGAGAAATTTGAAGCTCAACAATATAGTTGAAGTGGAG 1277
Db 709 TGGCTGCTGAGTACCTTCAGAGAAATTTGAAGCTCAACAATATAGTTGAAGTGGAG 768
Qy 1278 AAGCAATGGGTCTTCCTCCATCTTCGCGAGGCCAAATTACACACGAGACAGAGCTA 1337
Db 769 AAGCAATGGATTTGGAAGCAGTTAGAGGAATACGCCAACAGTACCACATGACATG 828
Qy 1338 AGAAGTAAATGAGAGAGCCCTAGATTTCCAGGAGCTGCCATTTAGGAAACGAAATGAAG 1397
Db 829 AAGCAATTTAGAAAGAGATGGG-----GAGAGAACCCAGAGGAGAACT 871
Qy 1398 GAACAGGATATTGGAAGCAGTTAGAGGAATACGCCAACAGTACCTCAATGACATGAAA 1457
Db 872 CAATAATAGTCATAAAACCTATTGTTGTAAGAGAGTAACTTCCCTGCTCCATGAGATG 931
Qy 1458 GAAATTTAGAAAGATGGGAGAGAACCCAGAGGACATTTGAAAAGACTTTGAAACAAATG 1517
Db 932 CATCTGAGGAGAGACCTGTG-----CAGGACATTTGAAAAGACTTTGAAACAAATG 984
Qy 1518 AGGCTTCAGACACAAAGGAGTAAATTCAGACAGAAATATAAGCTAGAGAGGGG 1577
Db 985 AGGCTTCAGACACAAAGGAGTAAATTCAGACAGAAATATAAGCTAGAGAGGGG 1044
Qy 1578 GTAAAAATTTGAAATTAATTTAGACAAATGATTTCTGATGAAAACATCTCTCCAGAGGAA 1637
Db 1045 GTAAAAATTTGAAATTAATTTAGACAAATGATTTCTGATGAAAACATCTCTCCAGAGGAA 1104
Qy 1638 GAGGCAATGATATACCAATGAACTTTGACCTTTCAGGATGCGATGAAATTAAGGAA 1697
Db 1105 GAGGCAATGATATACCAATGAACTTTGACCTTTCAGGATGCGATGAAATTAAGGAA 1164
Qy 1698 TATCAATGTTGTAAGGAGCAGTGGAGATTATACAGACAAAGCAATTTGAAAACCTTCACTGC 1757
Db 1165 TATCAATGTTGTAAGGAGCAGTGGAGATTATACAGACAAAGCAATTTGAAAACCTTCACTGC 1224
Qy 1758 CCAGAGCAGCATTTACAGAACTGACT 1784
Db 1225 CCAGAGCAGCGGTTTTCACGACGACT 1251
```

```
RESULT 15
BC068201 1745 bp mRNA linear PRI 31-MAR-2004
LOCUS BC068201 Homo sapiens cdna clone IMAGE:4771714, partial cds.
DEFINITION BC068201 Homo sapiens
ACCESSION BC068201.1 GI:45872598
VERSION BC068201.1
KEYWORDS Homo sapiens (human)
SOURCE Homo sapiens
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 1745)
AUTHORS Strausberg,R.L., Feingold,E.A., Grouse,L.H., Derge,J.G.,
Klausner,R.D., Collins,F.S., Wagner,L., Shenmen,C.M., Schuler,G.D.,
```

Altschul,S.F., Zeeberg,B., Buetow,K.H., Schaefer,C.F., Bhat,N.K., Hopkins,R.F., Jordan,H., Moore,T., Max,S.I., Wang,J., Hsieh,F., Diatchenko,L., Marusina,K., Farmer,A.A., Rubin,G.M., Hong,L., Stapleton,M., Soares,M.B., Bonaldo,M.F., Casavant,T.L., Scheetz,T.E., Brownstein,M.J., Ustin,T.B., Toshiyuki,S., Carninci,P., Prange,C., Raha,S.S., Loquellano,N.A., Peters,G.J., Abramson,R.D., Mullany,S.J., Bosak,S.A., McEwan,P.J., McKernan,K.J., Malek,J.A., Gunaratne,P.H., Richards,S., Wexley,K.C., Hale,S., Garcia,A.M., Gay,L.J., Hulyk,S.W., Villalon,D.K., Muzny,D.M., Sodergren,E.J., Lu,X., Gibbs,R.A., Fahey,J., Helton,E., Kettman,M., Madan,A., Rodriguez,S., Sanchez,A., Whiting,M., Madan,A., Young,A.C., Shevchenko,Y., Bouffard,G.G., Blakesley,R.W., Touchman,J.W., Green,E.D., Dickson,M.C., Rodriguez,A.C., Grimwood,J., Schmutz,E., Myers,R.M., Butterfield,Y.S., Krzywinski,M.I., Skalska,U., Smalhus,D.E., Schnerch,A., Schein,J.E., Jones,S.J. and Marra,M.A.

Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences

Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)

JOURNAL
PUBMED
REFERENCE
AUTHORS
JOURNAL

REMARK
COMMENT

NIH-MGC Project URL: <http://mgc.nci.nih.gov>
Contact: MGC help desk
Email: cgabbs-remail.nih.gov
Tissue Procurement: ATCC
cDNA Library Preparation: CLONTECH Laboratories, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Genome Sequence Centre,
BC Cancer Agency, Vancouver, BC, Canada
info@bcgsc.bc.ca

Steve Jones, Sarah Barber, Mabel Brown-John, Yaron Butterfield, Andy Chan, Steve S. Chand, William Chow, Alison Cloutier, Ruth Featherstone, Malachi Griffith, Obi Griffith, Ran Guin, Nancy Liao, Kim MacDonald, Amara Masson, Mike R. Mayo, Josh Moran, Ryan Morin, Teika Olson, Diana Palmquist, Anca Petrescu, Anna Liisa Prabhu, Parvaneh Saeedi, JR Santos, Angélique Schnerch, Ursula Skalska, Duane Smalhus, Jeff Stott, Miranda Tsai, George Yang, Jacquie Schein, Asim Siddiqui, Rob Holt, Marco Marra.

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
Series: IRAL Plate: 51 Row: k Column: 4
This clone was selected for full length sequencing because it passed the following selection criteria: Hexamer frequency ORF analysis.

FEATURES
source

1. 1745
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:4771714"
/tissue type="Testis, embryonal carcinoma"
/clone lib="NIH MGC_61"
/lab host="DH10B"
/note="Vector: pDNR-LIB"

ORIGIN

Query Match 15.6%; Score 369.4; DB 9; Length 1745;
Best Local Similarity 66.0%; Pred. No. 1.5e-75;
Matches 551; Conservative 0; Mismatches 281; Indels 3; Gaps 1;

Qy 9 ACCATGATAGTACGATGCTGATTAAGGCCATCGGCAAGTGCCTTCGGGAAGCATAC 68
Db 536 ATCATGGAGAGATGTTAGTACTACAGAGATTGGAGAGGTTTCATTGGAAAGCCATT 595
Qy 69 TTAGCTAAAGGGAATCATAGATGCAAGCCTGTGTCTATAAAGAGATCAATTTGAAAG 128

Db 596 CTGTGTTAAATCTACAGAAGATGGCAGACAGATATGTTATCAAGGAAATTAACATCTCAAGA 655
QY 129 ATGCCCATACAAGAAAAAGAGCTTTCAAGAAAGAGTATCTCTCGGAAAAAGATGAAA 188
Db 656 ATGTCAGTAAGAAAAAGAGAAATCAAGGAGAGAGATGTCAGTATTCGCAACATGAAG 715
QY 189 CATCCCAACATTTGTAGCCCTTCTTCAATTCATTTCAAGGAAATGGCAGGCTGTTTATTGTA 248
Db 716 CATCAAAATATTTGTCAGATATAGAGAACATTTGAAGAAATGGCTCTCTCTACATAGTA 775
QY 249 ATGGAATATTTGATGGAGGGATCTCATGAAAAGATCAATAGACAAACGGGGTGTGTTA 308
Db 776 ATGGATTACTGTGAGGGAGGGATCTGTTTAAGCGAATAAATGCTCAGAAAGGCGTTTTG 835
QY 309 TTTTACTGAGATCAGATCCTCGGTTGGTTTGTACAGATTTCTTAGGACTAAACATATT 368
Db 836 TTTCAAGAGGATCAGATTTGGACTGGTTTGTACAGATATGTTGGCCCTGAAACATGTA 895
QY 369 CATGACAGGAAGATATTACACAGGGACATAAAAGCTCAGAACATTTTCTTAGCAAGAAC 428
Db 896 CATGATAGAAAATTTCTTCATCGACACATTAATCTCAGAACATATTTTAACTAAAGAT 955
QY 429 GGAATGGTGGGAAAGCTTTGGGACCTTTGGTATAGCAAGAGTCCTGAATAATTCATGGAA 488
Db 956 GGAACAGT---ACAACTTGGAGATTTGGAATTTGTAATTTGCTAGAGTTCTTAATAGTACTGTAGAG 1012
QY 489 CTTGCTCGAACTTGTATTGGAACACCTTACTACCTGTCGCCAGAGATCTGTCAAGATAAA 548
Db 1013 CTGGCTCGAACTTGTATAGGAGCCCNATACCTTGTCACTGAAATCTGTGAAACAAA 1072
QY 549 CCCTACAAACAATAAAACGGATATTTGGTCTCTTGGCTGTGTCTTATATGAGCTCTGCACA 608
Db 1073 CCTTACAATAATAAAAGTGACATTTGGGCTCTGGGGTGTGCTTATGAGCTGTGTACA 1132
QY 609 CTTAAACATCCTTTTGGGGTAACAACTTACAGACAGCTGGTTCTGAAGATTTGTCAAGCA 668
Db 1133 CTTAAACATGCTTTTGAAGCTGGCAGTATGAAAAACCTGGTACTGAAGATAATATCTGGA 1192
QY 669 CATTTTGGCCCAATATCTCCGGGGTTTTCTCTGTGAGCTCCATTCCTTGATATCTCAGCTC 728
Db 1193 TCTTTTCCACCTGTGTCTTTGCAATTTATCTCTATGATCTCCGAGTTTGGTGTCTCAGTTA 1252
QY 729 TTTCAAGTATCTCCTCGAGACCGACCATCCATAAATTCATTTTGAAAAGGCCCTTTTATA 788
Db 1253 TTTAAAAGAAATCCTTAGGGATAGACCATCAGTCAACTCCATATTGGAGAAAGGTTTATA 1312
QY 789 GAGAACTTTATTCCAAATATTTGACTCTCTGAGGTCAATTCAGGAAGAAATTCAGTC 843
Db 1313 GCCAAACGCATTGAAAAGTTTCTCTCTCTCAGCTTATTCAGAGAGAAATTTTGTC 1367

Search completed: November 8, 2004, 00:08:20
Job time : 9758 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 7, 2004, 19:19:58 ; Search time 1060 Seconds
(without alignments)
11736.913 Million cell updates/sec

Title: US-10-730-010-1
Perfect score: 2370
Sequence: 1 cattggagaccatgataag.....ctgtcatgccgttacgagcg 2370

Scoring table: IDENTITY_NUC
Gapop 10_0 , Gapext 1.0

Searched: 4134886 seqs, 2624710521 residues
Total number of hits satisfying chosen parameters: 8269772

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-Processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N Geneseq_23Sep04:*

- 1: geneseqn1980s:*
- 2: geneseqn1990s:*
- 3: geneseqn2000s:*
- 4: geneseqn2001as:*
- 5: geneseqn2001bs:*
- 6: geneseqn2002as:*
- 7: geneseqn2002bs:*
- 8: geneseqn2003as:*
- 9: geneseqn2003bs:*
- 10: geneseqn2003cs:*
- 11: geneseqn2003ds:*
- 12: geneseqn2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB ID	Description
1	2370	100.0	2370	8	ABSS7149 Human cDN
2	1665	70.3	2240	6	ABK12426 Novel hum
3	1665	70.3	2816	12	ADJ96571 Human Nim
4	1656.2	69.9	1965	6	ABK12425 cDNA enco
5	1654	69.8	2052	6	ABK12424 cDNA enco
6	1449.4	61.2	1947	4	AAS06745 Polyucle
7	1349	56.9	1918	8	ABZ59717 Human NFK
8	1338	56.5	1896	8	ABZ59716 Human NOV
9	1199.6	50.6	1587	8	ABX72263 Human NOV
10	1176.8	49.7	1453	8	ABX72262 Human NOV
11	602	25.4	1632	10	ABZ77151 Human pro
12	597.4	25.2	1434	10	ADF44502 Mouse kin
13	597.4	25.2	2407	10	ADF44489 Mouse kin
14	369.4	15.6	4683	8	ABK13160 Human cDN
15	369.4	15.6	5426	5	AAS11582 Human cDN
16	369.4	15.6	5448	4	AAL58367 Human pol
17	369.4	15.6	5448	5	ADQ98576 DNA enco
18	369.4	15.6	5448	9	ADB48336 Novel hum
19	369.4	15.6	5508	8	AAL51591 Human ser
20	369.4	15.6	5532	4	AAL58366 Human pol
21	369.4	15.6	5532	5	ADQ98575 DNA enco

22	369.4	15.6	5532	9	ADB48335
23	369.4	15.6	5583	12	ADJ96572
24	368	15.5	3645	5	AAS11558 Human cDN
25	368	15.5	3729	8	AAL51590 Human ser
c	361.4	15.2	5514	4	AAI60152 Human pol
c	361.4	15.2	5514	4	AAI60153 Human pol
27	358.8	15.1	2343	6	ABV75946 Ser/Thr/T
28	358.8	15.0	4263	3	AAA09328 Human can
29	355.8	15.1	1588	8	ABX72261 Human NOV
30	262.2	11.1	1588	8	ABX72261 Human NOV
31	262.2	11.1	1588	12	ADH42550 Novel hum
32	260.6	11.0	1781	8	ABX72260 Human NOV
33	260.6	11.0	1781	12	ADH42548 Novel hum
34	260.6	11.0	2257	8	ABX72259 Human NOV
35	260.6	11.0	2257	12	ADH42546 Novel hum
36	254.6	10.7	1578	8	ABZ68773 Nucleotid
37	254.2	10.7	1581	8	ABZ68773 Nucleotid
38	254.2	10.7	1991	12	ADI40889 Human kin
39	254.2	10.7	2050	10	ADC99119 Human KPP
40	254.2	10.7	2102	6	ABZ69800 Human ser
41	254.2	10.7	2110	6	AAZ38850 Human kin
42	254.2	10.7	2220	4	AAK51477 Human pol
43	254.2	10.7	2326	12	ADJ96573 Human Nim
44	254.2	10.7	3255	12	ADM16424 DNA enco
45	253.8	10.7	876	4	AAH06097 Human cDN

ALIGNMENTS

RESULT 1
ABSS7149
ID ABSS7149 standard; cDNA; 2370 BP.
XX AC ABSS7149;
XX DT 04-FEB-2003 (first entry)
XX DE Human cDNA encoding a serine/threonine kinase.
XX KW Human; ss; gene; serine/threonine kinase; lung carcinoma; chromosome 13.
XX OS Homo sapiens.
XX FH Key Location/Qualifiers
XX FT CDS 12..1934
XX FT /*tag= a
XX FT /product= "Ser/Thr kinase"
XX PN WO200281727-A2.
XX PD 17-OCT-2002.
XX PF 02-APR-2002; 2002WO-US010156.
XX PR 03-APR-2001; 2001US-00824583.
XX PR 05-JUN-2001; 2001US-00873404.
XX PA (PEKE) PE CORP NY.
XX PI Webster M, Yan C, Di Francesco V, Beasley EM;
XX DR WPI; 2003-058562/05.
XX DR P-PSDB; ABG72000.
XX PT Novel human kinase protein expressed in lung carcinoma and placenta is
XX PT useful to diagnose and treat diseases and disorders associated with
XX PT expression or activity of the protein.
XX PS Claim 22; Fig 1A; 101pp; English.
XX CC The invention relates to an isolated human kinase peptide of the
XX CC serine/threonine, an allelic variant or orthologue, and encoded by a
XX CC nucleic acid that hybridises under stringent conditions with the cDNA and

CC gene sequences appearing as ABS57149 and ABS57150, or a fragment
CC comprising at least 10 contiguous amino acids. Also included are an
CC antibody that selectively binds to the kinase, a gene chip comprising the
CC nucleic acids, a transgenic non human animal comprising the nucleic
CC acids, a nucleic acid vector comprising the nucleic acids, a host cell-
CC containing the vector and expressing the kinase, identifying a
CC modulator/binding agent of the kinase (comprising contacting the peptide
CC with an agent and determining if the agent has modulated function,
CC expression or activity of the peptide or formed a complex with it). The
CC molecules of the invention are useful to diagnose and treat a disorder
CC characterised by aberrant expression of the protein (e.g. lung
CC carcinoma). Agents which modulate the function or activity of the protein
CC are useful to treat diseases or disorders mediated by human kinase
CC protein. The gene for encoding the novel kinase is located on chromosome
CC 13. The present sequence is the cDNA encoding the novel human kinase
XX

SQ Sequence 2370 BP; 796 A; 456 C; 550 G; 568 T; 0 U; 0 Other;

	Query Match	100.0%	Score 2370;	DB 8;	Length 2370;
	Best Local Similarity	100.0%;	Pred. No. 0;		
	Matches 2370;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
Qy	1	CATTGGAGACCATGTAAGTACGATGTGATTAAAGCCATCGGCCAAGTGCTTCGGGA	60		
Db	1	CATTGGAGACCATGTAAGTACGATGTGATTAAAGCCATCGGCCAAGTGCTTCGGGA	60		
Qy	61	AAGCATCTTAGCTTAAGGGAAATCAGATAGCAGCACTGTGTCATAAAGAGATCAATT	120		
Db	61	AAGCATCTTAGCTTAAGGGAAATCAGATAGCAGCACTGTGTCATAAAGAGATCAATT	120		
Qy	121	TTGAAAAGATGCCCATCAAGAAAAGAGCTTCAAGAAAAGAGTGAATCTTCTGAAA	180		
Db	121	TTGAAAAGATGCCCATCAAGAAAAGAGCTTCAAGAAAAGAGTGAATCTTCTGAAA	180		
Qy	181	AGATGAACATCCCAACATTTGATGAGCTTCTCAATTCATTCAGAAATGGCAGCTGT	240		
Db	181	AGATGAACATCCCAACATTTGATGAGCTTCTCAATTCATTCAGAAATGGCAGCTGT	240		
Qy	241	TTATTGTAATGGAATTTGATGAGGGGATCTCATGAAAAGATCAATAGACAACGGG	300		
Db	241	TTATTGTAATGGAATTTGATGAGGGGATCTCATGAAAAGATCAATAGACAACGGG	300		
Qy	301	GTGTGTTATTAGTGAAGATCAGATCTCGTGGTGGTTGTACAGATTTCTTAGGACTAA	360		
Db	301	GTGTGTTATTAGTGAAGATCAGATCTCGTGGTGGTTGTACAGATTTCTTAGGACTAA	360		
Qy	361	AACATATTTCATGACAGAGATATTACACAGGGACATATAAGCTCAGAACATTTTCTTA	420		
Db	361	AACATATTTCATGACAGAGATATTACACAGGGACATATAAGCTCAGAACATTTTCTTA	420		
Qy	421	GCAAGAACGGAAATGGTGGCAAGCTTCGGGACCTTTGGTATAGCAAGAGTCTTGAATAT	480		
Db	421	GCAAGAACGGAAATGGTGGCAAGCTTCGGGACCTTTGGTATAGCAAGAGTCTTGAATAT	480		
Qy	481	CCATGGAATGCTCGAACTTGATTCGAACACTTACTACTCTGCTCCAGAGATCTGTC	540		
Db	481	CCATGGAATGCTCGAACTTGATTCGAACACTTACTACTCTGCTCCAGAGATCTGTC	540		
Qy	541	AGATAAACCCCTACAACAATAAAACGGATTTGGTCTCTTGGCTGTGCTTATATGAGC	600		
Db	541	AGATAAACCCCTACAACAATAAAACGGATTTGGTCTCTTGGCTGTGCTTATATGAGC	600		
Qy	601	TCTGCACTTAACATCTTTTGGGGTAAACACTTACAGAGCTGGTCTGAAATTT	660		
Db	601	TCTGCACTTAACATCTTTTGGGGTAAACACTTACAGAGCTGGTCTGAAATTT	660		
Qy	661	GTCAAGCACATTTTGGCCCAATATCTCCGGGTTTCTCGTGAGCTCCATTCCTTGATAT	720		
Db	661	GTCAAGCACATTTTGGCCCAATATCTCCGGGTTTCTCGTGAGCTCCATTCCTTGATAT	720		
Qy	721	CTAGCTCTTTCAAGTATCTCTCGAGACCGACCATCAATAAATTCATTTTGAAGAAGC	780		
Db	721	CTAGCTCTTTCAAGTATCTCTCGAGACCGACCATCAATAAATTCATTTTGAAGAAGC	780		

Qy	781	CTTTTTAGAGAAATCTTATCCCAATATTTGACTCTCTGAGGTCAATTCAGGAAGATTC	840
Db	781	CTTTTTAGAGAAATCTTATCCCAATATTTGACTCTCTGAGGTCAATTCAGGAAGATTC	840
Qy	841	GTCAATGCTTATATGACAGAGCGCCAGCTTCTCGACATGCTGGGAAGTGGTCC	900
Db	841	GTCAATGCTTATATGACAGAGCGCCAGCTTCTCGACATGCTGGGAAGTGGTCC	900
Qy	901	AGAAGTGTAAATACAAAAGTGAATTCGGGGAAAGTGCACCAAGATCAAGGATAT	960
Db	901	AGAAGTGTAAATACAAAAGTGAATTCGGGGAAAGTGCACCAAGATCAAGGATAT	960
Qy	961	CTGTGCCAATTAAGAGGAATGCTATATTCATAGAAATGAATGGAGACCACCTGGAG	1020
Db	961	CTGTGCCAATTAAGAGGAATGCTATATTCATAGAAATGAATGGAGACCACCTGGAG	1020
Qy	1021	CCGAGAGCCAGATCTTATAAAATGATAGAAAGACCCCAAAATTTGCTGCTGTCTGGAC	1080
Db	1021	CCGAGAGCCAGATCTTATAAAATGATAGAAAGACCCCAAAATTTGCTGCTGTCTGGAC	1080
Qy	1081	ATTATGATTTATTTATGCTCAACTGATGCTGAGGAGAGAGCCCAAAACCAAGTT	1140
Db	1081	ATTATGATTTATTTATGCTCAACTGATGCTGAGGAGAGAGCCCAAAACCAAGTT	1140
Qy	1141	ATCACCTTATTCCTCAAGAAAATCTGGAGTTGAGGATTCAGGTAACGAGGAGCATG	1200
Db	1141	ATCACCTTATTCCTCAAGAAAATCTGGAGTTGAGGATTCAGGTAACGAGGAGCATG	1200
Qy	1201	GTCCATCCCAAGTCAATGGCTGCTGAGTACCTTCAGAGAAAATTTGAAGCTCAACAAT	1260
Db	1201	GTCCATCCCAAGTCAATGGCTGCTGAGTACCTTCAGAGAAAATTTGAAGCTCAACAAT	1260
Qy	1261	ATAAGTTCAGAGTGGAGAGCAATTTGGTCTTCTGCTCATCTTCTGCGAGCCCAATTACA	1320
Db	1261	ATAAGTTCAGAGTGGAGAGCAATTTGGTCTTCTGCTCATCTTCTGCGAGCCCAATTACA	1320
Qy	1321	ACCAGAGCAAGAGCTTAAGAAAGTAAATGGAGAAAGCCCTAGATTCAGAGAGCTGCATTTA	1380
Db	1321	ACCAGAGCAAGAGCTTAAGAAAGTAAATGGAGAAAGCCCTAGATTCAGAGAGCTGCATTTA	1380
Qy	1381	GGAAAAAGAAATGAAGAACAGGAATTTGGAGAGCTTAGAGGAAATACGCCAAAGT	1440
Db	1381	GGAAAAAGAAATGAAGAACAGGAATTTGGAGAGCTTAGAGGAAATACGCCAAAGT	1440
Qy	1441	ACCTCAATGACATGAAAGAAATTAAGAAAGAGATGGGAGAGAACAGAGACATTTGAAA	1500
Db	1441	ACCTCAATGACATGAAAGAAATTAAGAAAGAGATGGGAGAGAACAGAGACATTTGAAA	1500
Qy	1501	AAGACTTGAACAAATAGAGCTTCAGAACACAAAGGAAAGTAAATAATCCAGAACAGAAAT	1560
Db	1501	AAGACTTGAACAAATAGAGCTTCAGAACACAAAGGAAAGTAAATAATCCAGAACAGAAAT	1560
Qy	1561	ATAAGCTTAAGAGGGGTAAATTTGAAATTAATTTAGACAAATGATTTCTGATGAAA	1620
Db	1561	ATAAGCTTAAGAGGGGTAAATTTGAAATTAATTTAGACAAATGATTTCTGATGAAA	1620
Qy	1621	ACATCTCCCAAGAGGAGCAATGGATATACCAATGAAACTTTGACCTTTGAGGATG	1680
Db	1621	ACATCTCCCAAGAGGAGCAATGGATATACCAATGAAACTTTGACCTTTGAGGATG	1680
Qy	1681	GCATGAGTTTAAAGGAATATGATGTGTAAGGAGCATGGAGATTATACAGACAAAGCAT	1740
Db	1681	GCATGAGTTTAAAGGAATATGATGTGTAAGGAGCATGGAGATTATACAGACAAAGCAT	1740
Qy	1741	TTGAAAAATCTTCACTGCCAGAGCAGCATTTACAGAACTGACCTGGCTCAGTTTCTCT	1800
Db	1741	TTGAAAAATCTTCACTGCCAGAGCAGCATTTACAGAACTGACCTGGCTCAGTTTCTCT	1800
Qy	1801	TCCTGGATATCTCTGCTCATTTCTTCTGGAAATCTCCATTCAGAGGAGCATCTTA	1860
Db	1801	TCCTGGATATCTCTGCTCATTTCTTCTGGAAATCTCCATTCAGAGGAGCATCTTA	1860


```
QY 841 GTCATGCTTATATGACAGCAGGAGCGCCAGCTTCTCGACATGCTGGGAAGTGTGCC 900
Db 934 GTCATGCTTATATGACAGCAGGAGCGCCAGCTTCTCGACATGCTGGGAAGTGTGCC 993
QY 901 AGAAGTGTAAATACAAAAGTGAAGATTCGGGGAAAGTGCCCAACCAAGATCAAGATAT 960
Db 994 AGAAGTGTAAATACAAAAGTGAAGATTCAGGGAAGTGGCCACCAAGATCAAGATAT 1053
QY 961 CTGTGCCAATTAAAGGAATGCTATATTTGCTATAGAAATGAATGAGACCAACAGCTGAG 1020
Db 1054 CTGTGCCAATTAAAGGAATGCTATATTTGCTATAGAAATGAATGAGACCAACAGCTGAG 1113
QY 1021 CCCAAGGCCAGATCTATAAAATGATAGAAAGACCCAAATTTGCTGTCTGTGGAC 1080
Db 1114 CCCAAGGCCAGATCTATAAAATGATAGAAAGACCCAAATTTGCTGTCTGTGGAC 1173
QY 1081 ATTATGATTATATTATCTCAACTTGATGTGATGAGAGAGAGCCCAAAACCAAGTT 1140
Db 1174 ATTATGATTATATTATCTCAACTTGATGTGATGAGAGAGAGCCCAAAACCAAGTT 1233
QY 1141 ATCACCTTATTCCTCAAGAAATATCTGAGTTGAGGATTAACGTCAGGAACAGGCGATG 1200
Db 1234 ATCACCTTATTCCTCAAGAAATATCTGAGTTGAGGATTAACGTCAGGAACAGGCGATG 1293
QY 1201 GTCCATCCCCAAGTCAATGGCTGCTGAGTAGTACCTTCAGAGAAAATTTGAAGCTCAACAAT 1260
Db 1294 GTCCATCCCCAAGTCAATGGCTGCTGAGTAGTACCTTCAGAGAAAATTTGAAGCTCAACAAT 1353
QY 1261 ATAGTTGAAAGTGAGAGCAATTTGGTCTTCCTGCTCCATCTTCGCGAGCCCAATTTACA 1320
Db 1354 ATAGTTGAAAGTGAGAGCAATTTGGTCTTCCTGCTCCATCTTCGCGAGCCCAATTTACA 1413
QY 1321 ACCAGAGCAAGAGCTAAGAAATTAAGAAAGATGGGAGAGCTAGATTCAGGAGCTGCATTTA 1380
Db 1414 ACCAGAGCAAGAGCTAAGAAATTAAGAAAGATGGGAGAGCTAGATTCAGGAGCTGCATTTA 1473
QY 1381 GGAAGAACGAATGAGGAACAGAAATATGGAAGCAGTTAGAGGAATAGCCACAGT 1440
Db 1474 GGAAGAACGAATGAGGAACAGAAATATGGAAGCAGTTAGAGGAATAGCCACAGT 1533
QY 1441 ACCTCAATGACATGAAGAAATTAAGAAAGATGGGAGAGCAACCAAG----- 1488
Db 1534 ACCAATGACATGAAGAAATTAAGAAAGATGGGAGAGCAACCAAGACTCAA 1593
QY 1489 ----- 1488
Db 1594 AAATAAGTCATAAAACCTATTGTTGTAAGAGAGTAACCTGCTGTCATCAAGATGCAT 1653
QY 1489 -----AGCACATTGAAAGACTTGAACAAATCAGGCTTCAGA 1527
Db 1554 CTGAGGGAAGACCTGTGAGAGCAATTTGAAAGACTTGAACAAATCAGGCTTCAGA 1713
QY 1528 ACACAAAGGAAAGTAAATCCAGAACAGAAATATAAGCTTAAGAGGGGGTAAATTTG 1587
Db 1714 ACACAAAGGAAAGTAAATCCAGAACAGAAATATAAGCTTAAGAGGGGGTAAATTTG 1773
QY 1588 AAATTAATTTAGCAATGATTTCTGATGAAACATCCTCCAGAGAGAGAGGCAATGG 1647
Db 1774 AAATTAATTTAGCAATGATTTCTGATGAAACATCCTCCAGAGAGAGAGGCAATGG 1833
QY 1648 ATATACCAAAATGAACTTTGACCTTTGAGGATGCGATGAAGTTTAAGGAATATGAATGTG 1707
Db 1834 ATATACCAAAATGAACTTTGACCTTTGAGGATGCGATGAAGTTTAAGGAATATGAATGTG 1893
QY 1708 TAAAGGAGCATGGAGATTTATACAGAAAGCAATTTGAAACACTTCACTGCCAGAGAGCAG 1767
Db 1894 TAAAGGAGCATGGAGATTTATACAGAAAGCAATTTGAAACACTTCACTGCCAGAGAGCAG 1953
QY 1768 CATTTACAGAACTGACT 1784
Db 1954 GGTTTTCCACCGACT 1970
```

```
RESULT 3
ADJ96571
ID ADJ96571 standard; DNA; 2816 BP.
XX
AC ADJ96571;
XX
DT 06-MAY-2004 (first entry)
XX
DE Human Nim-A related protein kinase NEK5 DNA SeqID 28.
XX
Gene; ds; kinase; human; tyrosine protein kinase;
KW serine/threonine protein kinase; PTK; STK; gene therapy; cancer;
KW immune-related disease; cardiovascular disease; brain;
KW neuronal associated disease; metabolic; inflammatory disorder;
KW cytosolic; neuroprotective; immunomodulator; antiinflammatory;
KW Nim-A related protein kinase; NEK5.
XX
OS Homo sapiens.
OS 52.
XX WO2004006838-A2.
XX
PD 22-JAN-2004.
XX
XX 15-JUL-2003; 2003WO-US021730.
XX
XX 15-JUL-2002; 2002US-0395632P.
XX
PA (SUGB-) SUGEN INC.
XX
PI Whyte D, Manning G, Caenepeel S;
XX
DR WPI; 2004-122753/12.
XX
P-PSDB; ADJ96637.
XX
PT New nucleic acid molecule encoding a kinase polypeptide, useful for
PT preparing a composition for treating diseases or disorders, e.g., cancer,
PT or neurological, immunological or inflammatory disorders.
XX
PS Example 1; SEQ ID NO 28; 366pp; English.
XX
CC This invention relates to a novel isolated, enriched or purified nucleic
CC acid molecule that encodes a kinase polypeptide. Specifically, it relates
CC to human tyrosine and serine/threonine protein kinases (PTK's and STK's),
CC as well as protein kinase-like enzymes. The present invention describes
CC screening methods to identify agonists, antagonists and antibodies that
CC can be used to modulate the activity or function of the mammalian kinase
CC enzymes. As such, these compositions can be used for gene therapy
CC purposes to treat diseases or disorders including cancer, immune-related
CC diseases, cardiovascular disease, brain or neuronal associated disease,
CC metabolic and inflammatory disorders. Accordingly, they exhibit
CC cytosolic, neuroprotective, immunomodulator and antiinflammatory
CC activities. This polynucleotide sequence is a human kinase DNA sequence
CC of the invention.
XX
SQ Sequence 2816 BP; 977 A; 508 C; 675 G; 656 T; 0 U; 0 Other;
Query Match 70.3%; Score 1665; DB 12; Length 2816;
Best Local Similarity 94.5%; Pred. No. 0;
Matches 1774; Conservative 0; Mismatches 10; Indels 93; Gaps 1;
QY 1 CATTGGAGACCATGGATTAAGTACGATGTGATTAAGGCCATCGGCAAGGTGCTTCGGGA 60
Db 307 CATTGGAGACCATGGATTAAGTACGATGTGATTAAGGCCATCGGCAAGGTGCTTCGGGA 366
QY 61 AAGCATCTAGCTAAAGGAAATCAGATACGAGCACTGTGTCTATAAAGAGATCAATT 120
Db 367 AAGCATCTAGCTAAAGGAAATCAGATACGAGCACTGTGTCTATAAAGAGATCAATT 426
QY 121 TTGAAAAGATGCCCATACAGAAAAGAAAGCTTCAAGAAAAGAGTATTTCTTCGAAA 180
Db 427 TTGAAAAGATGCCCATACAGAAAAGAAAGCTTCAAGAAAAGAGTATTTCTTCGAAA 486
```

181 AGATGAACATCCCAACATTTAGCCTTCTCAATTCATTTCAAGAGAAATGCGAGCTGT 240
Db
487 AGATGAACATCCCAACATTTAGCCTTCTCAATTCATTTCAAGAGAAATGCGAGCTGT 546
Qy
241 TTATTGTAATGGAATATTGTGATGGAGGGGATCTCATGAAAGGATCAATAGACAAACGGG 300
Db
547 TTATTGTAATGGAATATTGTGATGGAGGGGATCTCATGAAAGGATCAATAGACAAACGGG 606
Qy
301 GTGTGTTATTAGTGAAGATCAGATCCTCGGTGGTTGTACAGATTTCTCTPAGACTAA 360
Db
607 GTGTGTTATTAGTGAAGATCAGATCCTCGGTGGTTGTACAGATTTCTCTPAGACTAA 666
Qy
361 ACATATTGATGACGAGAGATATTACACAGGACATAAAAGCTCAGAACATTTTCTTTA 420
Db
667 AACATATTGATGACGAGAGATATTACACAGGACATAAAAGCTCAGAACATTTTCTTTA 726
Qy
421 GCAAGAACGGAATGGTGGCAAGCTTTGGGAGCTTTGGTATACAGAGTCCTGATATAT 480
Db
727 GCAAGAACGGAATGGTGGCAAGCTTTGGGAGCTTTGGTATACAGAGTCCTGATATAT 786
Qy
481 CCATGGAACTTGTCTGAACTTGTATTGGAACACCTTACTACCTGTCCCCAGAGATCTGTC 540
Db
787 CCATGGAACTTGTCTGAACTTGTATTGGAACACCTTACTACCTGTCCCCAGAGATCTGTC 846
Qy
541 AGAATAAACCCCTACACAAATAAACGGATATTGGTCTCTTGGCTGTCTTATATGAGC 600
Db
847 AGAATAAACCCCTACACAAATAAACGGATATTGGTCTCTTGGCTGTCTTATATGAGC 906
Qy
601 TCTGCACACTTAAACATCTTTGAGGGTAACAACTTACAGAGCTGGTCTTGAAGATTT 660
Db
907 TCTGCACACTTAAACATCTTTGAGGGTAACAACTTACAGAGCTGGTCTTGAAGATTT 966
Qy
661 GTACAGCACATTTGGCCCAATATCTTCGGGGTTTCTCGTGAGCTCCATCTCTTGATAT 720
Db
967 GTACAGCACATTTGGCCCAATATCTTCGGGGTTTCTCGTGAGCTCCATCTCTTGATAT 1026
Qy
721 CTCAGCTCTTTCAAGTATCTCTCGAGACCGACCATCCATAAATTCATTTGAAAGGC 780
Db
1027 CTCAGCTCTTTCAAGTATCTCTCGAGACCGACCATCCATAAATTCATTTGAAAGGC 1086
Qy
781 CTTTTTTAGAGATCTTATCCCAATATTTGACCTCTCGAGTCAATCAGGAAGAAATCA 840
Db
1087 CTTTTTTAGAGATCTTATCCCAATATTTGACCTCTCGAGTCAATCAGGAAGAAATCA 1146
Qy
841 GTCAATGCTTATATGACAGACGAGCGCGAGCTTCTCGACATGCTGGGAGGTGTC 900
Db
1147 GTCAATGCTTATATGACAGACGAGCGCGAGCTTCTCGACATGCTGGGAGGTGTC 1206
Qy
901 AGAAGTGTAATAACAAAAGTGAGATTCGGGGGAAAGTGCCCAACCAAGATCAAGGATAT 960
Db
1207 AGAAGTGTAATAACAAAAGTGAGATTCAGGGGAAAGTGCCCAACCAAGATCAAGGATAT 1266
Qy
961 GTGTGCGCAATTAAGAGGATGCTATATGCTATAGAAATGAATGGAGACCAACAGCTGAG 1020
Db
1267 GTGTGCGCAATTAAGAGGATGCTATATGCTATAGAAATGAATGGAGACCAACAGCTGAG 1326
Qy
1021 CCCAGAAGCCAGATCTATAAAATGATAGAAAGACCCAAATTCCTCTGCTGTGGAC 1080
Db
1327 CCCAGAAGCCAGATCTATAAAATGATAGAAAGACCCAAATTCCTCTGCTGTGGAC 1386
Qy
1081 ATTATGATPATTATTATGCTCAACTTGATPACTGCTGAGGAGAGACCCCAAAACCAAGTT 1140
Db
1387 ATTATGATPATTATTATGCTCAACTTGATPACTGCTGAGGAGAGACCCCAAAACCAAGTT 1446
Qy
1141 ATCACCCCTATTCTCAAGGAATATCTGGAGTTGAGGATTCGGTCAGGAACGAGGCATG 1200
Db
1447 ATCACCCCTATTCTCAAGGAATATCTGGAGTTGAGGATTCGGTCAGGAACGAGGCATG 1506
Qy
1201 GTCCATCCCAAGTCAATGGCTGCTGAGTACTCTTTCAGAGAAAATTTGAAGCTCAACAAT 1260
Db
1507 GTCCATCCCAAGTCAATGGCTGCTGAGTACTCTTTCAGAGAAAATTTGAAGCTCAACAAT 1566
Qy
1261 ATAAGTTGAAAGTGGAAGCAATTTGGGTCTTCGTCCATCTCTGCGCGAGCAATTTACA 1320

1567 ATAAAGTTGAAAGTGGAAGCAATTTGGGTCTTCGTCCATCTCTGCGCGAGCAATTTACA 1626
Qy
1321 ACCAGACACAGAGCTTAAGAGTAATGGAGAGAGCTAGATTCAGAGAGCTGCGCATTTA 1380
Db
1527 ACCAGACACAGAGCTTAAGAGTAATGGAGAGAGCTAGATTCAGAGAGCTGCGCATTTA 1686
Qy
1381 GGAAGAACGAAATGAAGGAAACAGAAATATTGGAGAGCAGTTAGAGGAAATACGCCAACAGT 1440
Db
1687 GGAAGAACGAAATGAAGGAAACAGAAATATTGGAGAGCAGTTAGAGGAAATACGCCAACAGT 1746
Qy
1441 ACCTCAATGACATGAAGAAATATTAGAAAGAGATGGGGAGAGAACCCAG----- 1488
Db
1747 ACCACATGACATGAAGAAATATTAGAAAGAGATGGGGAGAGAACCCAGAGAGAACTCAA 1806
Qy
1489 ----- 1488
Db
1807 AAATAAGTCAATAAACCTATTTTGGTGAAGAGAGTAACCTGCTGTCCTCAATCAAGATGCAT 1866
Qy
1489 -----AGGACATTTGAAAGAGACTTTGAAACAAATGAGGCTTTGAGA 1527
Db
1867 CTGAGGAGAGAGCACCTGTGACGACATTTGAAAGAGACTTTGAAACAAATGAGGCTTTGAGA 1926
Qy
1528 ACACAAAGGAAAGTAAAGATCCAGACAGAAATATAAGCTAAGAGGAGGGGTAAGATTTG 1587
Db
1927 ACACAAAGGAAAGTAAAGATCCAGACAGAAATATAAGCTAAGAGGAGGGGTAAGATTTG 1986
Qy
1588 AAATAATTTAGACAAATGTATTTCTGATCAAAACATCTCTCCAGAGAGAGGCAATGG 1647
Db
1987 AAATAATTTAGACAAATGTATTTCTGATCAAAACATCTCTCCAGAGAGAGGCAATGG 2046
Qy
1648 ATATACCAATGAAACTTTGACCTTTGAGGATGGCATGAAATTTAAGGAATATGAATGTG 1707
Db
2047 ATATACCAATGAAACTTTGACCTTTGAGGATGGCATGAAATTTAAGGAATATGAATGTG 2106
Qy
1708 TAAAGGAGCAGTGGAGATTTATACAGACAAAGCATTTGAAAGAACTTCACTGCCAGAGAGCAG 1767
Db
2107 TAAAGGAGCAGTGGAGATTTATACAGACAAAGCATTTGAAAGAACTTCACTGCCAGAGAGCAG 2166
Qy
1768 CATTTACAGAACTGACT 1784
Db
2167 GGTTCCTCCAGCACT 2183
RESULT 4
ABK12425
ID ABK12425 standard; cDNA; 1965 BP.
XX
AC ABK12425;
XX
DT 18-JUN-2002 (first entry)
XX
DE cDNA encoding novel human kinase protein #2.
XX
KW Novel human protein; NHP; serine-threonine kinase; brain;
KW calcium/calmodulin-dependent kinase; myosin light chain kinase;
KW biological disorder; spleen; placenta; chromosome 6; gene; ss.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT CDS 1..1965
FT /*tag= a
FT /product= "Human kinase protein #2"
XX
FN WO200218555-A2.
XX
PD 07-MAR-2002.
XX
PF 28-AUG-2001; 2001WO-US026776.
XX
PR 31-AUG-2000; 2000US-0229280P.
XX

PA (LEXI-) LEXICON GENETICS INC.
XX Friddle CU, Hilbun E, Nepomnichy B, Hu Y;
XX WPI; 2002-292200/33.
XX P-PSDB; AAU77929.
XX Novel polynucleotide encoding novel human protein sharing structural
XX similarity with animal kinases e.g. serine-threonine, calcium/calmodulin-
XX dependent, and myosin light chain kinases, useful as probes and primers.
XX Disclosure; Page 39-40; 46pp; English.
XX The present invention relates to the isolation of novel human proteins
XX (NHPs) and the polynucleotide sequences encoding them. The NHPs of the
XX invention are kinase proteins and share structural similarity to serine-
XX threonine, calcium/calmodulin-dependent, and myosin light chain kinases.
XX The sequences of the invention are useful for treating biological
XX disorders. The polynucleotide sequences encoding the kinase proteins can
XX be used as primers and probes. The sequences are also useful for
XX identifying mutations associated with a particular disease and also in a
XX prognostic or diagnostic assay. The present sequence encodes human
XX protein kinase #2 which is expressed in a broad range of human tissues
XX such as brain, spleen, and placenta. The gene encoding protein kinase #2
XX is located on chromosome 6
XX Sequence 1965 BP; 658 A; 377 C; 458 G; 472 T; 0 U; 0 Other;
XX
XX Query Match 69.9%; Score 1656.2; DB 6; Length 1965;
XX Best Local Similarity 96.1%; Pred. No. 0;
XX Matches 1710; Conservative 0; Mismatches 63; Indels 6; Gaps 1;
XX
XX 12 ATGATAGTACGATGTGATTAAAGCCATCGGCAAGTGCCTTCGGGAAGCATACTTA 71
XX Db 1 ATGATAGTACGATGTGATTAAAGCCATCGGCAAGTGCCTTCGGGAAGCATACTTA 60
XX
XX 72 GCTAAAGGGAATCAGATAGCAAGCATGTGTCTATAAAGAGATCAATTTGAAAGATG 131
XX Db 61 GCTAAAGGGAATCAGATAGCAAGCATGTGTCTATAAAGAGATCAATTTGAAAGATG 120
XX
XX 132 CCCATACAGAAAGAGAGCTTCAAGAAAGATGATCTCTCGAAGAGATGAACAT 191
XX Db 121 CCCATACAGAAAGAGAGCTTCAAGAAAGATGATCTCTCGAAGAGATGAACAT 180
XX
XX 192 CCCAACATTTGATGCTTCTTCAATTCATTTCAAGAGAAATGGCAGGCTGTTTATGTAATG 251
XX Db 181 CCCAACATTTGATGCTTCTTCAATTCATTTCAAGAGAAATGGCAGGCTGTTTATGTAATG 240
XX
XX 252 GAATATTGTGATGAGGGGATCTCATGAAGAGATCAATGACACACGGGCTGTGTTATTT 311
XX Db 241 GAATATTGTGATGAGGGGATCTCATGAAGAGATCAATGACACACGGGCTGTGTTATTT 300
XX
XX 312 AGTGAAGATCAGATCTCGGTGGTGGTGTACAGATTTCTTAGGACTTAAACATATTTTCAAT 371
XX Db 301 AGTGAAGATCAGATCTCGGTGGTGGTGTACAGATTTCTTAGGACTTAAACATATTTTCAAT 360
XX
XX 372 GACAGGAGATATTACACAGGACATATAAGCTCAGAACATTTTCTTAGCAAGACGGA 431
XX Db 361 GACAGGAGATATTACACAGGACATATAAGCTCAGAACATTTTCTTAGCAAGACGGA 420
XX
XX 432 ATGTGGCAAGCTTTGGGACCTTTGGTATAGCAAGAGTCTGTAATATTCATTTGGAACTT 491
XX Db 421 ATGTGGCAAGCTTTGGGACCTTTGGTATAGCAAGAGTCTGTAATATTCATTTGGAACTT 480
XX
XX 492 GCTCGAACTTGTATTGGAACACCTTACTACTGTCCAGAGATCTGTCTAGAAATAAACCC 551
XX Db 481 GCTCGAACTTGTATTGGAACACCTTACTACTGTCCAGAGATCTGTCTAGAAATAAACCC 540
XX
XX 552 TACAAACATTAACCGGATATTGGTCTCTTGGCTGTCTTATATGAGCTCTGCACACTT 611
XX Db 541 TACAAACATTAACCGGATATTGGTCTCTTGGCTGTCTTATATGAGCTCTGCACACTT 600
XX
XX 612 AAACATCTTTTGGGGTAAACACTTACAGCAGCTGGTCTTGAAGATTTTGTCAAGCACAT 671
XX
XX 601 AAACATCTTTTGGGGTAAACACTTACAGCAGCTGGTCTTGAAGATTTTGTCAAGCACAT 660
XX Db
XX 672 TTTGCCCCCAATATCTCCGGGGTTTTCTCGTGAGCTCCATCTCTTGATATCTCAGCTCTTT 731
XX Db
XX 661 TTTGCCCCCAATATCTCCGGGGTTTTCTCGTGAGCTCCATCTCTTGATATCTCAGCTCTTT 720
XX Db
XX 732 CAGATATCTCTCGAGACCGACCATCCATTAATTTCCATTTGAAAGGCCCTTTTAGAG 791
XX Db
XX 721 CAGATATCTCTCGAGACCGACCATCCATTAATTTCCATTTGAAAGGCCCTTTTAGAG 780
XX Db
XX 792 AATCTTATTTCCAAATATTTGACTCTCTGAGGTCAATTCAGGAAGAAATTCAGTCACATGCTT 851
XX Db
XX 781 AATCTTATTTCCAAATATTTGACTCTCTGAGGTCAATTCAGGAAGAAATTCAGTCACATGCTT 840
XX Db
XX 852 ATATCGAGCAGAGAGCCCGCAGCTTCTCGACATGCTGGAAGAGTGTGCCAGAAATGTAAA 911
XX Db
XX 841 ATATCGAGCAGAGAGCCCGCAGCTTCTCGACATGCTGGAAGAGTGTGCCAGAAATGTAAA 900
XX Db
XX 912 ATACAAAAGTGAGATTCCGGGGAAGTGCCCAACAGATCAAGCATATCTGTGCCAAT 971
XX Db
XX 901 ATACAAAAGTGAGATTCCAGGGAAGTGCCCAACAGATCAAGCATATCTGTGCCAAT 960
XX Db
XX 972 AAAAGGAATGCTATATTTGCATAGAAATGAATGGAGACACACAGCTGGAGGCCAGAGGCC 1031
XX Db
XX 961 AAAAGGAATGCTATATTTGCATAGAAATGAATGGAGACACACAGCTGGAGGCCAGAGGCC 1020
XX Db
XX 1032 AGATCTATAAATGATAGAAAGCCCAAAATTCGCTGCTGTGGACATTTATGATTAT 1091
XX Db
XX 1021 AGATCTATAAATGATAGAAAGCCCAAAATTCGCTGCTGTGGACATTTATGATTAT 1080
XX Db
XX 1092 TATTATGCTCAACTTGTATGCTGAGGAGAGGCCCAAAACCAAGTTATCAACCTTATT 1151
XX Db
XX 1081 TATTATGCTCAACTTGTATGCTGAGGAGAGGCCCAAAACCAAGTTATCAACCTTATT 1140
XX Db
XX 1152 CCTCAAGAAATATCTGGAGTTAGGATTTACGGTCAGGAAGAGGCGATGTCCTATCCCA 1211
XX Db
XX 1141 CCTCAAGAAATATCTGGAGTTAGGATTTACGGTCAGGAAGAGGCGATGTCCTATCCCA 1200
XX Db
XX 1212 AGTCAATGGCTGTGTAGTACCTTTCAGAGAAATTTGAAAGCTCAACATATAAGTTGAAA 1271
XX Db
XX 1201 AGTCAATGGCTGTGTAGTACCTTTCAGAGAAATTTGAAAGCTCAACATATAAGTTGAAA 1260
XX Db
XX 1272 GTGGAGAGCAATTTGGGTCTTCTGCTCCAGGACCAATTTACACAGAGACAA 1331
XX Db
XX 1261 GTGGAGAGCAATTTGGGTCTTCTGCTCCAGGACCAATTTACACAGAGACAA 1320
XX Db
XX 1332 GAGCTAAAGAAATTAATGGAGAGAGCGCTAGATTCCAGGAGCTGCCATTTAGGAAAACGAA 1391
XX Db
XX 1321 GAGCTAAAGAAATTAATGGAGAGAGCGCTAGATTCCAGGAGCTGCCATTTAGGAAAACGAA 1380
XX Db
XX 1392 ATGAAGGAACAGGAATATTGGAAGCAGTTAGAGAAATACGCCAACAGTACCTCAATGAC 1451
XX Db
XX 1381 ATGAAGGAACAGGAATATTGGAAGCAGTTAGAGAAATTAAGTCAATAAACCTATTGTTGTAAGAGAGTAAC 1440
XX Db
XX 1452 ATGAAGGAATTAAGAAAGATGGGAGAGAACCA-----GAGGACATTTGAAAAGAC 1505
XX Db
XX 1441 CTGCTGCTTCCATCAGATGTCATCTGAGGAGAGAGCCTGTGTCAGGACATTTGAAAAGAC 1500
XX Db
XX 1506 TTGAAACCAATGAGGCTTTCAGAACACAAAGGAAAGTAAATAATCCAGAACAGAAATATAAA 1565
XX Db
XX 1501 TTGAAACCAATGAGGCTTTCAGAACACAAAGGAAAGTAAATAATCCAGAACAGAAATATAAA 1560
XX Db
XX 1566 GCTAAAGAGGGGGTAAATTTGAAATTAATTTAGACAAATGATTTCTGTGAAAACATC 1625
XX Db
XX 1561 GCTAAAGAGGGGGTAAATTTGAAATTAATTTAGACAAATGATTTCTGTGAAAACATC 1620
XX Db
XX 1626 CTCCAAAGAGGAGGCAATGGATATACCAATGAAACTTTGACCTTTGAGGATGGCATG 1685
XX Db
XX 1621 CTCCAAAGAGGAGGCAATGGATATACCAATGAAACTTTGACCTTTGAGGATGGCATG 1680
XX Db
XX 1686 AAGTTTAAAGGAATATGATGTGTAAGAGAGCATGGAGATATACAGAACAGCAATTTGAA 1745
XX Db

Db 1681 AAGTTTAAGGATATGAATGTGTAAAGGAGCATGGAGATTATACAGACAAAGCAATTGAA 1740
Qy 1746 AAACCTTCACTGCCAGAGCAGCATTTTACAGAACTGACT 1784
Db 1741 AAACCTTCACTGCCAGAGCAGGCTTTTCCAGCAGACT 1779

RESULT 5
ABK12424
ID ABK12424 standard; cDNA; 2052 BP.
XX
AC ABK12424;
XX
DT 18-JUN-2002 (first entry)
XX
DE cDNA encoding novel human kinase protein #1.
XX

XX Novel human protein; NHP; serine-threonine kinase; brain;
KW calcium/calmodulin-dependent kinase; myosin light chain kinase;
KW biological disorder; spleen; placenta; chromosome 6; gene; ss.
XX
OS Homo sapiens.

XX Key Location/Qualifiers
FH 1. .2052
FT CDS /tag= a
FT /product= "human kinase protein #1"
XX

XX WO200218555-A2.
XX 07-MAR-2002.
XX 28-AUG-2001; 2001WO-US026776.
XX 31-AUG-2000; 2000US-0229280P.
XX (LEXI-) LEXICON GENETICS INC.
XX
XX Fridele CJ, Hilbun E, Nepomnichy B, Hu Y;
XX WPI: 2002-292200/33.
XX P-FSDB; AAU77928.

XX Novel polynucleotide encoding novel human protein sharing structural
PT similarity with animal kinases e.g. serine-threonine, calcium/calmodulin-
PT dependent, and myosin light chain kinases, useful as probes and primers.
XX
XX Claim 1; Page 37-38; 46pp; English.

XX The present invention relates to the isolation of novel human proteins
CC (NHPs) and the polynucleotide sequences encoding them. The NHPs of the
CC invention are kinase proteins and share structural similarity to serine-
CC theonine, calcium/calmodulin-dependent, and myosin light chain kinases.
CC The sequences of the invention are useful for treating biological
CC disorders. The polynucleotide sequences encoding the kinase proteins can
CC be used as primers and probes. The sequences are also useful for
CC identifying mutations associated with a particular disease and also in a
CC prognostic or diagnostic assay. The present sequence encodes human
CC protein kinase #1 which is expressed in a broad range of human tissues
CC such as brain, spleen, and placenta. The gene encoding protein kinase #1
CC is located on chromosome 6

XX Sequence 2052 BP; 698 A; 388 C; 482 G; 484 T; 0 U; 0 Other;
XX
XX Query Match 69.8%; Score 1654; DB 6; Length 2052;
XX Best Local Similarity 94.5%; Pred. No. 0;
XX Matches 1763; Conservative 0; Mismatches 10; Indels 93; Gaps 1;

Qy 12 ATGATAAGTACGATGTATTAAGCCATCGGCAAGTGCCTTCGGGAAAGCATCTTA 71
Db 1 ATGATAAGTACGATGTATTAAGCCATCGGCAAGTGCCTTCGGGAAAGCATCTTA 60
Qy 72 GCTAAAGGGAATCAGATAGCAAGCAGCTGTGTCTATAAAGAGATCAATTTTGAAGATG 131

Db 61 GCTAAAGGGAATCAGATAGCAAGCAGCTGTGTCTATAAAGAGATCAATTTTGAAGATG 120
Qy 132 CCATACAGAAAGAGAGCTTCAAGAAAGAGTGTATTTCTTGGGAAAGATGAACAT 191
Db 121 CCATACAGAAAGAGAGCTTCAAGAAAGAGTGTATTTCTTGGGAAAGATGAACAT 180
Qy 192 CCACAACTTGTAGCTTCTTCAATTTCAAGAGAAATGCGAGGCTGTTTATTGTAAAG 251
Db 181 CCACAACTTGTAGCTTCTTCAATTTCAAGAGAAATGCGAGGCTGTTTATTGTAAAG 240
Qy 252 GAATATTGTATGAGGAGGATCTCATGAAGAAGATCAATAGACAACGGGGTGTATTAT 311
Db 241 GAATATTGTATGAGGAGGATCTCATGAAGAAGATCAATAGACAACGGGGTGTATTAT 300
Qy 312 AGTGAAGATCAGATCTCTCGGTGTTGTACAGATTTCTTAGAGTCTATAAACAATTTT 371
Db 301 AGTGAAGATCAGATCTCTCGGTGTTGTACAGATTTCTTAGAGTCTATAAACAATTTT 360
Qy 372 GACAGGAAGATATTAACACAGGACATAAAGCTCAGAACATTTTCTTAGCAAGAACGGA 431
Db 361 GACAGGAAGATATTAACACAGGACATAAAGCTCAGAACATTTTCTTAGCAAGAACGGA 420
Qy 432 ATGTCGCAAGCTCTCGGACTTTGTATAGCAAGAGTCTGTAATTAATTCATGGAAC 491
Db 421 ATGTCGCAAGCTCTCGGACTTTGTATAGCAAGAGTCTGTAATTAATTCATGGAAC 480
Qy 492 GCTCGAACTTGTATGGAACACCTTACTACTCTCCCGAGAGATCTGTGAGATAAACC 551
Db 481 GCTCGAACTTGTATGGAACACCTTACTACTCTCCCGAGAGATCTGTGAGATAAACC 540
Qy 552 TACAACATAAAGCGGATATTTGGTCTCTTGGCTGTCTTATATAGCTCTGACACTT 611
Db 541 TACAACATAAAGCGGATATTTGGTCTCTTGGCTGTCTTATATAGCTCTGACACTT 600
Qy 612 AAACATCTTCTTGGAGGTAAACACTTACAGCAGCTGTTCTGAAGATTTGTCAAGCACAT 671
Db 601 AAACATCTTCTTGGAGGTAAACACTTACAGCAGCTGTTCTGAAGATTTGTCAAGCACAT 660
Qy 672 TTGCCCCCAATATCTCCGGGTTTCTGAGAGTCCATCTCTTGAATATCTCAGCTCTTT 731
Db 661 TTGCCCCCAATATCTCCGGGTTTCTGAGTCCATCTCTTGAATATCTCAGCTCTTT 720
Qy 732 CAAGTATCTCTCGAGACCGACCATCCATAAATTCATTTTGAAGAGCCCTTTTAGAG 791
Db 721 CAAGTATCTCTCGAGACCGACCATCCATAAATTCATTTTGAAGAGCCCTTTTAGAG 780
Qy 792 AATCTTATTTCCAAATATTTGACTCTCTGAGTCAATTCAGAAAGATTCAGTCAATGCTT 851
Db 781 AATCTTATTTCCAAATATTTGACTCTCTGAGTCAATTCAGAAAGATTCAGTCAATGCTT 840
Qy 852 ATATGAGAGCAGGAGCGCCAGCTTCTGACATGCTGGGAAGGTGCTCAGAAAGTATAA 911
Db 841 ATATGAGAGCAGGAGCGCCAGCTTCTGACATGCTGGGAAGGTGCTCAGAAAGTATAA 900
Qy 912 ATACAAAAAGTACAGATTCCTGGGAAAGTCCCAAGATCAAGGATATCTGTGCCAAT 971
Db 901 ATACAAAAAGTACAGATTCCTGGGAAAGTCCCAAGATCAAGGATATCTGTGCCAAT 960
Qy 972 AAAGGAATGCTATATTGATGAATGAATGAATGAATGAATGAATGAATGAATGAATGA 1031
Db 961 AAAGGAATGCTATATTGATGAATGAATGAATGAATGAATGAATGAATGAATGAATGA 1020
Qy 1032 AGATCTATATAAATGATAGAAAGACCCAAATTTGCTGTCTGTGGACATTTATGATAT 1091
Db 1021 AGATCTATATAAATGATAGAAAGACCCAAATTTGCTGTCTGTGGACATTTATGATAT 1080
Qy 1092 TATTATGCTCAACTGATATGCTGAGGAGGAGAGCCCAAAACCAAGTTATCACCTATT 1151
Db 1081 TATTATGCTCAACTGATATGCTGAGGAGGAGAGCCCAAAACCAAGTTATCACCTATT 1140
Qy 1152 CTTCAAGAAATACCTGAGTTTACGCTCAGGAACCGAGGATGCTCCATCCCA 1211

1141 CCTCAAGAAATACCTGGAGTTGAGGATTACGGTCAGGAAACGAGGCATGTCCTCCATCCCA 1200
 1212 AGTCATATGGCTCTGAGTACCTTCAGAGAAATTTGAAGCTCAACAATATAGTTGAAA 1271
 1201 AGTCAATGGCTCTGAGTACCTTCAGAGAAATTTGAAGCTCAACAATATAGTTGAAA 1260
 1272 GTGAGAGCAATGGGTCTTCCTCCATCTTCGCCGAGCCAAATTTACAAACGAGACAA 1331
 1261 GTGAGAGCAATGGGTCTTCCTCCATCTTCGCCGAGCCAAATTTACAAACGAGACAA 1320
 1332 GAGTAAAGTAAATGAGAGAGCTTCAGATTCAGAGAGCTGCATTTAGGAAACGAA 1391
 1321 GAGTAAAGTAAATGAGAGAGCTTCAGATTCAGAGAGCTGCATTTAGGAAACGAA 1380
 1392 ATGAAGCAACAGGAATATTGGAAGCAGTTAGAGGAATATCGCCAAACAGTACCTCAATGAC 1451
 1381 ATGAAGCAACAGGAATATTGGAAGCAGTTAGAGGAATATCGCCAAACAGTACCTCAATGAC 1440
 1452 ATGAAGCAATTTAGAAAGATGGGAGAGAACCGAG----- 1488
 1441 ATGAAGCAATTTAGAAAGATGGGAGAGAACCGAGAGAGAACTCAAAAAATAGTCAAT 1500
 1489 ----- 1488
 1501 AAAACCTATTGTTGTAAGAGAGTAACCTGCTGTCATCAGATGATCGATGAGGAGAA 1560
 1489 -----AGGACATGTAAGAGAGCTTGAACCAATAGAGGCTTCAGAACACAAAGGAA 1538
 1561 GCACCTGTGAGGAGCAATTTGAAGAGACTTGAACCAATAGAGGCTTCAGAACACAAAGGAA 1620
 1539 AGTAAATATCCAGACAGAAATATTAAGCTTAAGAGAGGCTTAAATTTGAATTTAATTTA 1598
 1621 AGTAAATATCCAGACAGAAATATTAAGCTTAAGAGAGGCTTAAATTTGAATTTAATTTA 1680
 1599 GACAAATGATTTCTGATGAAACATCTCCAGAGAGAGAGGCAATGGATATACCAAT 1658
 1681 GACAAATGATTTCTGATGAAACATCTCCAGAGAGAGAGGCAATGGATATACCAAT 1740
 1659 GAACTTTGACCTTTGAGGTGCGATGAGCTTTAAGGATATCAATGCTTAAGGAGCAT 1718
 1741 GAACTTTGACCTTTGAGGTGCGATGAGCTTTAAGGATATCAATGCTTTAAGGAGCAT 1800
 1719 GGAGATTATACAGACAAAGCATTTGAAAACTTCACTGCCAGAGAGAGCATTTACAGAA 1778
 1801 GGAGATTATACAGACAAAGCATTTGAAAACTTCACTGCCAGAGAGAGCATTTACAGAA 1860
 1779 CTGACT 1784
 1861 CAGACT 1866

RESULT 6

AAS06745
 ID AAS06745 standard; cDNA; 1947 BP.
 XX AC AAS06745;
 XX DT 12-SEP-2001 (first entry)
 XX DE Polynucleotide sequence encoding human protein kinase #45.
 XX KW Human; protein kinase; PTK; STK; cancer; cardiovascular disease;
 KW metabolic disorder; immune related disease; neurological disorder;
 KW neurodegenerative disorder; inflammatory disorder; infectious disease;
 KW reproductive disorder; gene therapy; ss.
 XX OS Homo sapiens.
 XX FN WO200138503-A2.
 XX PD 31-MAY-2001.
 XX PF 22-NOV-2000; 2000WO-US032085.

24-NOV-1999; 99US-0167482P.
 (SUGEN-) SUGEN INC.
 Plowman GP, Whyte D, Manning G, Sudarsanam S, Martinez R;
 Flanagan P, Clary D;
 WPI; 2001-343950/36.
 P-PSDB; AAU03545.
 Nucleic acids encoding human kinase polypeptides, useful for preventing
 diagnosing and/or treating e.g. cancer, immune, cardiovascular and
 neuronal-associated diseases, and microbial infections.
 Example 1; Fig 1; 433pp; English.
 AAS06701-AAS06757 encode for novel human protein kinases #1-57. The novel
 protein kinases have been identified as members of the tyrosine or
 serine/threonine kinase (PTK and STK) families. The polynucleotides
 encoding protein kinases and the polypeptides may be used in the
 prevention, diagnosis and treatment of diseases associated with
 inappropriate kinase expression. For example, they may be used to treat
 cancers (especially cancers of haematopoietic origin), cardiovascular
 disease (e.g. atherosclerosis), metabolic disorders (e.g. diabetes),
 immune related diseases (e.g. rheumatoid arthritis), neurological
 disorders (e.g. schizophrenia), neurodegenerative disorders (e.g.
 Parkinson's disease), inflammatory disorders (e.g. asthma), infectious
 disease (e.g. HIV) and reproductive disorders (e.g. infertility).
 Additionally, polynucleotides encoding protein kinases may be used for
 gene therapy and as DNA probes in diagnostic assays. The protein kinase
 polypeptides may be used as antigens in the production of antibodies
 against the protein kinases and in assays to identify modulators of
 protein kinase expression and activity

Sequence 1947 BP; 638 A; 357 C; 455 G; 497 T; 0 U; 0 Other;

XX	Query Match	61.2%;	Score 1449.4;	DB 4;	Length 1947;
PR	Best Local Similarity	90.4%;	Pred. No. 0;		
XX	Matches 1602;	Conservative 0;	Mismatches 96;	Indels 75;	Gaps 2;
XX	12	ATGATAAGTACGATGATTAAGGCCATCGGCAAGTGCCTTCGGGAAGCATATCTTA	71		
XX	1	ATGGATAAGTACGATGATTAAGGCCATCGGCAAGTGCCTTCGGGAAGCATATCTTA	60		
XX	72	GCTAAAGGGAATCAGATAGCAAGCACTGTGTCTATAAAGAGATCAATTTTGAAGATG	131		
XX	61	GCTAAAGGGAATCAGATAGCAAGCACTGTGTCTATAAAGAGATCAATTTTGAAGATG	120		
XX	132	CCCATACAGAAAAAGAGCTTCAAGAAAGAGATGATTTCTTGGAAAAAGATGAAACAT	191		
XX	121	CCCATACAGAAAAAGAGCTTCAAGAAAGAGATGATTTCTTGGAAAAAGATGAAACAT	180		
XX	192	CCCACATTTGAGCTTCTTCAATTTCAAGAGAAATGCGAGCTGTTTATTGTAATG	251		
XX	181	CCCACATTTGAGCTTCTTCAATTTCAAGAGAAATGCGAGCTGTTTATTGTAATG	240		
XX	252	GAATATTGTGATGGAGGGGATCTCATGAAAAAGGATCAATAGACAAACGGGTGTTTATT	311		
XX	241	GAATATTGTGATGGAGGGGATCTCATGAAAAAGGATCAATAGACAAACGGGTGTTTATT	300		
XX	312	AGTGAAGATCAGATCTCGGTGTTGTTGATAGATTTCTTAGGACTTAAACATATTTCAT	371		
XX	301	AGTGAAGATCAGATCTCGGTGTTGTTGATAGATTTCTTAGGACTTAAACATATTTCAT	360		
XX	372	CACAGGAAGATATTACACAGGGACATAAAGCTCAGAACATTTTCTTAGCAAGAACGGA	431		
XX	361	CACAGGAAGATATTACACAGGGACATAAAGCTCAGAACATTTTCTTAGCAAGAACGGA	420		
XX	432	ATGGTGGCAAGCTTTGGGACTTTGGTATAGCAGAGCTCCTGATATATTCATGGAATTT	491		
XX	421	ATGGTGGCAAGCTTTGGGACTTTGGTATAGCAGAGCTCCTGATATATTCATGGAATTT	480		

QY 492 GCTGAACTTGTATTGGAACACCTTACTACTCTGCTCCCGAGAGATCTGTGAGATTAACCC 551
DB 481 GCTGAACTTGTATTGGAACACCTTACTACTCTGCTCCCGAGAGATCTGTGAGATTAACCC 540
QY 552 TACAACAATAAAGCGGATTTGGTCTCTTGGCTGTGTCTTATATAGCTCTGCACTT 611
DB 541 TACAACAATAAAGCGGATTTGGTCTCTTGGCTGTGTCTTATATAGCTCTGCACTT 600
QY 612 AAACATCTTTTTCAGGATTAACACTTACAGCAGCTGGTCTGAGATTTGTCAAGACAT 671
DB 601 AAACATCTTTTTCAGGATTAACACTTACAGCAGCTGGTCTGAGATTTGTCAAGACAT 660
QY 672 TTTGCCCAATATCTCCGGGTTTCTCGTGAGCTCCATTCCTTGATATCTCAGCTCTTT 731
DB 661 TTTGCCCAATATCTCCGGGTTTCTCGTGAGCTCCATTCCTTGATATCTCAGCTCTTT 720
QY 732 CAGATATCTCTCGAGACCGACCATCATTAATTCATTTGAAAAGGCCCTTTTAGAG 791
DB 721 CAGATATCTCTCGAGACCGACCATCATTAATTCATTTGAAAAGGCCCTTTTAGAG 780
QY 792 AATCTTATTTCCCAATATTTGACTCTCTGAGGTCAATTCAGAGAAATTCAGTCACATGCTT 851
DB 781 AATCTTATTTCCCAATATTTGACTCTCTGAGGTCAATTCAGAGAAATTCAGTCACATGCTT 821
QY 852 ATATGAGAGAGCGGCGGAGCTTCTCGACATGCTGGGAAGGTGCTCCAGAAAGTAA 911
DB 822 -----GGTGACTGTGTTGGATTTTGGCAGAGATTTTGGTGTGCGAGTCTTG 867
QY 912 ATACAAAAGAGTGAATTCGGGGAGAGTCCCAACCAAGATCAAGATATCTGCGCAAT 971
DB 868 ACACGTGTGTTGGTGTGAGTGTGAGCCACTGCACCCAGCTGTATATGTTTTTAA 927
QY 972 AAAAGGAATGCTATATTTGATAGAAATGAATGGAGACCAACAGCTGGAGCCCAAGGCC 1031
DB 928 AACATCCCTCTGTTTTCTT----- 946
QY 1032 AGATCTATAAATATGATAGAGACCCAAATTTGCTGTCTGTGGACATTTATGATAT 1091
DB 949 ----CAGATAAATATGATAGAGACCCAAATTTGCTGTCTGTGGACATTTATGATAT 1005
QY 1092 TATTTATGCTCACTTGTATGCTGAGGAGAGAGCCCAACCAAGTTATCACCTATT 1151
DB 1006 TATTTATGCTCACTTGTATGCTGAGGAGAGAGCCCAACCAAGTTATCACCTATT 1065
QY 1152 CCTCAAGAAATACTGGAGTTGAGGATTAACGTCAGGAAACGAGGATGTCCTCCCA 1211
DB 1066 CCTCAAGAAATACTGGAGTTGAGGATTAACGTCAGGAAACGAGGATGTCCTCCCA 1125
QY 1212 AGTCAATGCTGCTGAGTACCTTCAGAGAAATTTGAAGCTCAACAATATAGTTGAA 1271
DB 1126 AGTCAATGCTGCTGAGTACCTTCAGAGAAATTTGAAGCTCAACAATATAGTTGAA 1185
QY 1272 GTGAGAAACAATTTGGTCTTCTGTCATCTTCTGCGAGCCAAATTAACAACAGAGACA 1331
DB 1186 GTGAGAAACAATTTGGTCTTCTGTCATCTTCTGCGAGCCAAATTAACAACAGAGACA 1245
QY 1332 GAGCTAAGAGTAATGGAAGAGCTAGATTCAGAGAGCTGCCATTTAGGAAAACGAA 1391
DB 1246 GAGCTAAGAGTAATGGAAGAGCTAGATTCAGAGAGCTGCCATTTAGGAAAACGAA 1305
QY 1392 ATCAAGGAACAGGAATTTGGAAGCAGTTAGAGGAATACGCCAACAGTACCTCAATGAC 1451
DB 1306 ATCAAGGAACAGGAATTTGGAAGCAGTTAGAGGAATACGCCAACAGTACCAATGAC 1365
QY 1452 ATGAAGAAATTTAGAAAGAGATGGGAGAGAACAGAGGACATTTGAAAAGACTTGAA 1511
DB 1366 ATGAAGAAATTTAGAAAGAGATGGGAGAGAACAGAGGACATTTGAAAAGACTTGAA 1425
QY 1512 CAAATGAGCTTCAGAACCAAGGAAGTAAATTCAGAGACAGAAATATAAGCTTAAG 1571
DB 1426 CAAATGAGCTTCAGAACCAAGGAAGTAAATTCAGAGACAGAAATATAAGCTTAAG 1485
QY 1572 AAGGGGTTAAATTTGAAATTAATTTAGACAAATGTAATTTCTGTGAAAACATCTCCAA 1631

DB 1486 AAGGGGTTAAATTTGAAATTAATTTAGACAAATGTAATTTCTGTGAAAACATCTCCAA 1545
QY 1632 GAGGAAGAGGCAATGCGATATACCAATGAACTTTTGACCTTTGAGGATGGCATGAAGTTT 1691
DB 1546 GAGGAAGAGGCAATGCGATATACCAATGAACTTTTGACCTTTGAGGATGGCATGAAGTTT 1605
QY 1692 AAGGAATATGAATGTGTAAAGGAGCATGGAGATTTATACAGCAAAAGCATTTGAAAACCTT 1751
DB 1606 AAGGAATATGAATGTGTAAAGGAGCATGGAGATTTATACAGCAAAAGCATTTGAAAACCTT 1665
QY 1752 CACTGCCCAGAGCAGCATTTTACAGAACTGACT 1784
DB 1666 CACTGCCCAGAGCAGGGTTTTCACGAGACT 1698

RESULT 7

ABZ59717

ID ABZ59717 standard; DNA; 1918 BP.

XX AC ABZ59717;

XX DT 17-APR-2003 (first entry)

XX DE Human NEK-like serine/threonine protein kinase DNA # SEQ ID 4.

XX KW Human; NEK-like serine/threonine protein kinase; cytostatic; cardiant;
XX KW antiinflammatory; nootropic; neuroprotective; cancer; colon cancer;
XX KW cardiovascular disorder; diabetes; COPD; CNS disorder; ds.

XX OS Homo sapiens.

XX PN WO2003000903-A2.

XX PD 03-JAN-2003.

XX PF 24-JUN-2002; 2002WO-EP006948.

XX PR 25-JUN-2001; 2001US-0300068P.

XX PR 07-DEC-2001; 2001US-0336704P.

XX PA (FARB) BAYER AG.

XX PI Xiao Y;

XX DR WPI; 2003-184051/18.

XX PT New polynucleotide encoding a NEK-like serine/threonine kinase
XX PT polypeptide useful for treating diseases associated with kinase
XX PT dysfunction, e.g. cardiovascular disorders, cancer such as colon cancer,
XX PT diabetes and CNS disorders.

XX PS Disclosure; Fig 4; 149pp; English.

XX PS The invention relates to a newly isolated polynucleotide encoding an NEK-
XX CC like serine/threonine protein kinase. The activity of the polynucleotide
XX CC and polypeptide of the invention may be described as cytostatic,
XX CC cardiant, antiinflammatory, nootropic and neuroprotective. The expression
XX CC vector and reagent of the invention are useful for the preparation of a
XX CC medicament for modulating the activity of an NEK-like serine/threonine
XX CC kinase in a disease, such as cancer (e.g. colon cancer), cardiovascular
XX CC disorder, diabetes, COPD or CNS disorder. The polypeptides may also be
XX CC used to identify compounds which may act as activators or inhibitors at
XX CC the enzyme's active site, to raise specific antibodies which can block
XX CC the enzyme and effectively reduce its activity, as a bait protein in a
XX CC two-hybrid or three-hybrid assay to identify other proteins which bind to
XX CC or interact with the human NEK-like serine/threonine kinase polypeptide
XX CC and modulate its activity, and for the immunisation of mammals. The
XX CC current sequence represents human NEK-like serine/threonine protein
XX CC kinase encoding DNA

XX SQ Sequence 1918 BP; 615 A; 367 C; 460 G; 476 T; 0 U; 0 Other;

Query Match		56.9%; Score 1349; DB 8; Length 1918;
Best Local Similarity		87.6%; Pred. No. 0;
Matches 1564; Conservative		0; Mismatches 90; Indels 131; Gaps 3;
QY	1	CATTGGAGACCATGGATGAAGTACCATGTGATTAAAGGCCATCGGCGCAAGGTGCTTCGGGA 60
DB	12	CATTGGAGACCATGGATGAAGTACCATGTGATTAAAGGCCATCGGCGCAAGGTGCTTCGGGA 71
QY	61	AAGCATCTTACCTAAAGGGAATCAGATAGCAAGCACTGTGTTCATATAAAGAGATCAATT 120
DB	72	AAGCATCTTACCTAAAGGGAATCAGATAGCAAGCACTGTGTTCATATAAAGAGATCAATT 131
QY	121	TTGAAAGATGCCATCAAGAAAGAGAGCTTCAAGAGAAAGAGTCAATTCTTCGAAA 180
DB	132	TTGAAAGATGCCATCAAGAAAGAGAGCTTCAAGAGAAAGAGTCAATTCTTCGAAA 191
QY	181	AGATGAACATCCCAACATTTAGCCTTCTCAATTCATTTCAAGAAATGGCGAGCTGT 240
DB	192	AGATGAACATCCCAACATTTAGCCTTCTCAATTCATTTCAAGAAATGGCGAGCTGT 251
QY	241	TTATTGTAATGGAATTTGATCGAGGGATCTCATGAAAGGATCAATAGCAACGGG 300
DB	252	TTATTGTAATGGAATTTGATCGAGGGATCTCATGAAAGGATCAATAGCAACGGG 311
QY	301	CTGTGTTATTAGTCAAGATCAGATCCCTCGGTTGGTTTGTACAGATTTCTTAGGACTAA 360
DB	312	CTGTGTTATTAGTCAAGATCAGATCCCTCGGTTGGTTTGTACAGATTTCTTAGGACTAA 371
QY	361	AACATATTCTACAGGAAAGATTTACACAGGACATATAAGCTCAGAACATTTTCTTAA 420
DB	372	AACATATTCTACAGGAAAGATTTACACAGGACATATAAGCTCAGAACATTTTCTTAA 431
QY	421	GCAAGAACGGAATGTGGCAAGCTTGGGAGCTTTGGTATAGCAAGAGTCTCTGAATAATT 480
DB	432	GCAAGAACGGAATGTGGCAAGCTTGGGAGCTTTGGTATAGCAAGAGTCTCTGAATAATT 491
QY	481	CCATGGAACTTCTGCAACTTTGATTTGGAACACCTTACTACCTGTCCCGCAGAGATCTGTC 540
DB	492	CCATGGAACTTCTGCAACTTTGATTTGGAACACCTTACTACCTGTCCCGCAGAGATCTGTC 551
QY	541	AGATTAACCTTACACATATAAGCGATTTTGGTCTCTTGGGCTGTCTTATATGAGC 600
DB	552	AGATTAACCTTACACATATAAGCGATTTTGGTCTCTTGGGCTGTCTTATATGAGC 611
QY	601	TCTGCACACTTTAAACATCTTTGAGGGTAAACAATTACAGCAGCTGGTTCTGAAGATT 660
DB	612	TCTGCACACTTTAAACATCTTTGAGGGTAAACAATTACAGCAGCTGGTTCTGAAGATT 671
QY	661	GTCAAGACATTTTGGCCCAATNTCCGGGGTTTCTCGTAGCTCCATTCCTTGATAT 720
DB	672	GTCAAGACATTTTGGCCCAATNTCCGGGGTTTCTCGTAGCTCCATTCCTTGATAT 731
QY	721	CTCAGCTCTTTCAAGTATCTCTCGAGACCGACCATCCATAAATTCATTTTGAAGAGC 780
DB	732	CTCAGCTCTTTCAAGTATCTCTCGAGACCGACCATCCATAAATTCATTTTGAAGAGC 791
QY	781	CCTTTTATAGAACTCTTATTCGAAATTTTGTACTCTCTGAGGTCATTCAGGAAGATTC 840
DB	792	CCTTTTATAGAACTCTTATTCGAAATTTTGTACTCTCTGAGGTCATTCAGGAAGATTC 851
QY	841	GTCAATGCTTATATGACAGAGAGGCGCAGCTTCTCGACATCTGCGGAAGTGTGTC 900
DB	852	GTCAATGCTTATATGACAGAGAGGCGCAGCTTCTCGACATCTGCGGAAGTGTGTC 911
QY	901	AGAGTGTAAATAACAAAAGTGAATTCGGGGGAAAGTGCACCAAGATCAAGATAT 960
DB	912	AGAGGCTACTGGTGTGAGGAGTGGGCTCTCAAGGCTTTGGGCAGC-----T 958
QY	961	CTGTGCCAATTAAGAGGATCTATTTGCTAGAGAAATGAATGGAGA-CCACCAAGCTGGA 1019
DB	959	CTGTCTCTGAGGCTTTGAGGCTTACAGCCCTTGGGCTGTCTCACAGGCTGTGTTGAG 1018
QY	1020	GCCAGAGAGGCGAGATCTATAAATAATGATAGAAAGACCCAAAATTTGCTGTCTGTGGA 1079
DB	1019	TGCTGCGGCTTTCCAGATATAAATGATAGAAAGACCCAAAATTTGCTGTCTGTGGA 1078
QY	1080	CATTATGATTTATTTATGCTCACTTCAATGATGCTGAGGAGAGAGCCCAAAACCAAGT 1139
DB	1079	CATTATGATTTATTTATGCTCACTTCAATGATGCTGAGGAGAGAGCCCAAAACCAAGT 1138
QY	1140	TATCACCTTATTTCTCAAGAAATTAATCTGAGTTGAGGATTAACGTCAGGAAACAGGAT 1199
DB	1139	TATCACCTTATTTCTCAAGAAATTAATCTGAGTTGAGGATTAACGTCAGGAAACAGGAT 1198
QY	1200	GGTCCATCCCAAGTCAATGGCTGCTGAGTACCTTCAGAGAAAATTTGAAGCTCAACAA 1259
DB	1199	GGTCCATCCCAAGTCAATGGCTGCTGAGTACCTTCAGAGAAAATTTGAAGCTCAACAA 1258
QY	1260	TATAAGTTGAAGTGGAGAGCAATTTGGTCTTCTGCTCATCTTCTGCGGACCAATTTAC 1319
DB	1259	TATAAGTTGAAGTGGAGAGCAATTTGGTCTTCTGCTCATCTTCTGCGGACCAATTTAC 1284
QY	1320	AACAGAGACAGAGCTAAGAAATTAATGGAGAGAGCCCTAGATTTCCAGAGCTGCCATTT 1379
DB	1285	----- 1284
QY	1380	AGGAAAACAGAAATCAAGGAACAGGAATTTGGAGCAGTTTAGAGAAATACGCCACAG 1439
DB	1285	-----GGATATTGGAGCAGTTTAGAGAAATACGCCACAG 1321
QY	1440	TACCTCAATGACATCAAGAAATTAAGAAAGAGATGGGAGAGAACCCAGAGAGCATTTGAA 1499
DB	1322	TACCAATGACATCAAGAAATTAAGAAAGAGATGGGAGAGAACCCAGAGAGCATTTGAA 1381
QY	1500	AAAGACTTGAACAAATGAGGCTTCAGAACCAAGAGAAAGTAAATAATCCAGAACAGAAA 1559
DB	1382	AAAGACTTGAACAAATGAGGCTTCAGAACCAAGAGAAAGTAAATAATCCAGAACAGAAA 1441
QY	1560	TATAAGCTTAAGAGGGGTAAATTTGAAATTTAATTTAGACAAATGATTTCTTGATGAA 1619
DB	1442	TATAAGCTTAAGAGGGGTAAATTTGAAATTTAATTTAGACAAATGATTTCTTGATGAA 1501
QY	1620	AAATCTCTCCAGAGAGAGCAATGATATACCAATGAACTTTGACCTTTGAGGAT 1679
DB	1502	AAATCTCTCCAGAGAGAGCAATGATATACCAATGAACTTTGACCTTTGAGGAT 1561
QY	1680	GGCATGAAGTTTAAAGATATGATGCTTAAGGAGCATGGAGTTATACAGACAGAGCA 1739
DB	1562	GGCATGAAGTTTAAAGATATGATGCTTAAGGAGCATGGAGTTATACAGACAGAGCA 1621
QY	1740	TTTGAAGAACTTCACTGCGCAGAGCAGCATTTTACAGAACTGACT 1784
DB	1622	TTTGAAGAACTTCACTGCGCAGAGCAGGGTTTTCCACGAGACT 1666
RESULT 8		
ABZ59716		
ID	ABZ59716	standard; DNA; 1896 BP.
AC	ABZ59716;	
XX	17-APR-2003	(first entry)
DT	Human NEK-like serine/threonine protein kinase DNA # SEQ ID 1.	
DE	Human; NEK-like serine/threonine protein kinase; cytosolic; cardiant;	
KW	antiflammatory; neutropic; neuroprotective; cancer; colon cancer;	
KW	cardiovascular disorder; diabetes; COPD; CNS disorder; gene; ds.	
XX	Homo sapiens.	
OS		
XX		
PH	Key	Location/Qualifiers
FT	1..1896	
FT	/*tag= a	
FT	/product= "NEK-like serine/threonine protein kinase"	
XX		

PN WO2003000903-A2.
XX
PD 03-JAN-2003.
XX
PF 24-JUN-2002; 2002WO-EP006948.
XX
PR 25-JUN-2001; 2001US-0300068P.
XX
PR 07-DEC-2001; 2001US-0336704P.
XX
PA (FARB) BAYER AG.
XX
PI Xiao Y;
XX
PI WPI; 2003-184051/18.
DR P-PSDB; ABP71710.
XX
DR
XX
PT New polynucleotide encoding a NEK-like serine/threonine kinase
PT polypeptide useful for treating diseases associated with kinase
PT dysfunction, e.g. cardiovascular disorders, cancer such as colon cancer,
PT diabetes and CNS disorders.
XX
PS Claim 1 b; Fig 1; 149pp; English.
XX
CC The invention relates to a newly isolated polynucleotide encoding an NEK-
CC like serine/threonine protein kinase. The activity of the polynucleotide
CC and polypeptide of the invention may be described as cytostatic,
CC cardiant, antiinflammatory, nootropic and neuroprotective. The expression
CC vector and reagent of the invention are useful for the preparation of a
CC medicament for modulating the activity of an NEK-like serine/threonine
CC kinase in a disease, such as cancer (e.g. colon cancer), cardiovascular
CC disorder, diabetes, COPD or CNS disorder. The polypeptides may also be
CC used to identify compounds which may act as activators or inhibitors at
CC the enzyme's active site, to raise specific antibodies which can block
CC the enzyme and effectively reduce its activity, as a bait protein in a
CC two-hybrid or three-hybrid assay to identify other proteins which bind to
CC or interact with the human NEK-like serine/threonine kinase polypeptide
CC and modulate its activity, and for the immunisation of mammals. The
CC current sequence represents the human NEK-like serine/threonine protein
CC kinase encoding DNA
XX
SQ Sequence 1896 BP; 608 A; 360 C; 455 G; 473 T; 0 U; 0 Other;

Query Match 56.5%; Score 1338; DB 8; Length 1896;
Best Local Similarity 87.5%; Pred. No. 0;
Matches 1553; Conservative 0; Mismatches 90; Indels 131; Gaps 3;

QY 12 ATGGATAAGTACGATGATGATTAAGCCATCGGGCAAGTGCCTCGGAAAGCATACATTA 71
DB 1 ATGGATAAGTACGATGATGATTAAGCCATCGGGCAAGTGCCTCGGAAAGCATACATTA 60
QY 72 GCTAAAGGGAATCAGATAGCAAGCACTGTGTCTATAAAAGAGATCAATTTGAAAAGATG 131
DB 61 GCTAAAGGGAATCAGATAGCAAGCACTGTGTCTATAAAAGAGATCAATTTGAAAAGATG 120
QY 132 CCATACAGAAAAGAAAGAGCTTCAAGAAAGAGATGATCTTCTGGAAGAGATGAACAT 191
DB 121 CCATACAGAAAAGAAAGAGCTTCAAGAAAGAGATGATCTTCTGGAAGAGATGAACAT 180
QY 192 CCCAATGTAGTACCTTCTTCAATTCATTTCAAGAGATGGCAGGCTGTTTATTGTATG 251
DB 181 CCCAATGTAGTACCTTCTTCAATTCATTTCAAGAGATGGCAGGCTGTTTATTGTATG 240
QY 252 GAATATGTAGTGGGGATCTCATGAAGAGATCAATAGACACGGGTGTGTATT 311
DB 241 GAATATGTAGTGGGGATCTCATGAAGAGATCAATAGACACGGGTGTGTATT 300
QY 312 AGTGAAGATCAGATCCTCGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 371
DB 301 AGTGAAGATCAGATCCTCGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 360
QY 372 GACAGAGATATTACACGGGACATATAAGCTCAGAACATTTTCTTAGCAAGACGGA 431
DB 361 GACAGAGATATTACACGGGACATATAAGCTCAGAACATTTTCTTAGCAAGACGGA 420

QY 432 ATGTTGCAAGAGCTTGGGACCTTTGGTATAGCAAGAGTCTCTGAATAAATTCATGGAACCTT 491
DB 421 ATGTTGCAAGAGCTTGGGACCTTTGGTATAGCAAGAGTCTCTGAATAAATTCATGGAACCTT 480
QY 492 GCTGAACTTGTATTGGAAACACCTTACTACTCTGCCCCAGAGATCTCTCAGAAATAAACCC 551
DB 481 GCTGAACTTGTATTGGAAACACCTTACTACTCTGCCCCAGAGATCTCTCAGAAATAAACCC 540
QY 552 TACAACAATAAAACGGGATATTTGGTCTCTTTGGTGTGTCTTATATAGAGCTCTGCACACTT 611
DB 541 TACAACAATAAAACGGGATATTTGGTCTCTTTGGTGTGTCTTATATAGAGCTCTGCACACTT 600
QY 612 AAACATCTTTTGGGCTTAACAATTTACAGAGCTGGTCTGAAGATTCTGCAAGCAGAT 671
DB 601 AAACATCTTTTGGGCTTAACAATTTACAGAGCTGGTCTGAAGATTCTGCAAGCAGAT 660
QY 672 TTTCCCCCAATATCTCCGGGGTCTTCTGCTGAGCTCCATTCCTTTGATATCTCAGCTCTTT 731
DB 661 TTTCCCCCAATATCTCCGGGGTCTTCTGCTGAGCTCCATTCCTTTGATATCTCAGCTCTTT 720
QY 732 CAAGTATCTCTCGAGACCGACCATCCATAAATTCATTTTGAAGGCCCTTTTAGAG 791
DB 721 CAAGTATCTCTCGAGACCGACCATCCATAAATTCATTTTGAAGGCCCTTTTAGAG 780
QY 792 AATCTTATTTCCCAATATTTGACTCTCTGAGGTCAATTCAGGAAGAAATTCAGTCAATGCTT 851
DB 781 AATCTTATTTCCCAATATTTGACTCTCTGAGGTCAATTCAGGAAGAAATTCAGTCAATGCTT 840
QY 852 ATATGACAGACGAGCGCCAGCTTTCTGACATCTCTGGGAAGTGTGTCAGAAAGTGTGAAA 911
DB 841 ATATGACAGACGAGCGCCAGCTTTCTCGACATGCTGGGAAGGTGGTCCAGAGGCATACT 900
QY 912 ATACAAAAGTGAATTTCCGGGAAAGTGCCTCAAGGCCCTTGGGAGC-----TCTGCTCTGAG 947
DB 901 GGTGTGAGGAGTGGGCTCTCAAGGCCCTTGGGAGC-----TCTGCTCTGAG 947
QY 972 AAAAGGAATGCTATATTGCTATAGAAATGAATGGAGACCAACAGCTGGAGGCCAAGAGGCC 1031
DB 948 GCTTTGACAGCTACAGCCCTCGGGTGTCTCTCACAGGCTGTGTGAGTGTCTGGGCT 1007
QY 1032 AGATCT-ATAAAAATGATAGAAAGACCCAAAATTTGCTGTCTGTGGACATTTATGATTA 1090
DB 1008 TTTCCAGATAAAAATGATAGAAAGACCCAAAATTTGCTGTCTGTGGACATTTATGATTA 1067
QY 1091 TTATTATGCTCAACTTGTATGCTGAGGAGAGGCCCAACAAACCAAGTTATCACCCCTAT 1150
DB 1068 TTATTATGCTCAACTTGTATGCTGAGGAGAGGCCCAACAAACCAAGTTATCACCCCTAT 1127
QY 1151 TCCTCAAGAAAATACCTGGAGTTACGCTCAGGAAACGAGGATGGTCCATCCCC 1210
DB 1128 TCCTCAAGAAAATACCTGGAGTTACGCTCAGGAAACGAGGATGGTCCATCCCC 1187
QY 1211 AAGTCAATGGCTGCTGAGTACCTTCAAGAAAATTTGAAGCTCAACAATATAAGTTGAA 1270
DB 1188 AAGTCAATGGCTGCTGAGTACCTTCAAGAAAATTTGAAGCTCAACAATATAAGTTGAA 1247
QY 1271 AGTGAAGAACAAATGGGCTCTTCTGCTCCATCTTCTGCGGAGCCAAATTAACAACAGAGACA 1330
DB 1248 AGTGAAGAACAAAT----- 1262
QY 1331 AGAGCTAAGAGTAATGGAGAGAGCCTAGATTCCAGGAGCTGCCATTTAGGAAAAACGA 1390
DB 1263 ----- 1262
QY 1391 AATGAAGAAACAGAAATATTGGAAGCAGTTAGAGGAAATACGCCAARCAAGTACCTCAATGA 1450
DB 1263 -----GGAATATTGGAAGCAGTTAGAGGAAATACGCCAARCAAGTACCTCAATGA 1310
QY 1451 CATGAAGAAATTTAGAAAGAGATGGGAGAGAAACAGAGGACATTTGAAAAGACTTGAA 1510
DB 1311 CATGAAGAAATTTAGAAAGAGATGGGAGAGAAACAGAGGACATTTGAAAAGACTTGAA 1370

QY 1511 ACAATGAGGCTTCAGAACACAGAGGAAAGTAAATATCCAGACAGAAATATAAGCTAA 1570
 DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 1371 ACAATGAGGCTTCAGAACACAGAGGAAAGTAAATATCCAGACAGAAATATAAGCTAA 1430
 QY 1571 GAAGGGGTAAATTTGAAATTTAATTTAGACAAATGTATTTCTGATGAAACATCTCTCCA 1630
 DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 1431 GAAGGGGTAAATTTGAAATTTAATTTAGACAAATGTATTTCTGATGAAACATCTCTCCA 1490
 QY 1631 AGAGGAAGGCAATGATATACCAATGAAACTTTGACCTTTGAGATGGCATGAAGTT 1690
 DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 1491 AGAGGAAGGCAATGATATACCAATGAAACTTTGACCTTTGAGATGGCATGAAGTT 1550
 QY 1691 TAAAGGAATATGAATGTGTAAGGAGGAGCATGGAGATTATACAGACAAAGCATTTGAAAAACT 1750
 DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 1551 TAAAGGAATATGAATGTGTAAGGAGGAGCATGGAGATTATACAGACAAAGCATTTGAAAAACT 1610
 QY 1751 TCACCTGCCAGAGAGCATTTTACAGAACTGACT 1784
 DB 1611 TCACCTGCCAGAGAGCGGTTTTCACAGCAGACT 1644

RESULT 9

ABX72263
 ID ABX72263 standard; cDNA; 1587 BP.

XX AC ABX72263;

XX DT 03-JUN-2003 (first entry)

XX DE Human NOVX polynucleotide #94.

XX Human; NOVX; gene; ss; metabolic disorder; cardiomyopathy; diabetes; ASD;
 KW hypertension; congenital heart defect; aortic stenosis; valve disease;
 KW atrial septal defect; atrioventricular canal defect; ductus arteriosus;
 KW pulmonary stenosis; subaortic stenosis; ventricular septal defect; VSD;
 KW tuberosus sclerosis; scleroderma; atherosclerosis; infectious disease;
 KW obesity; anorexia; neurodegenerative disorder; Alzheimer's disease;
 KW Parkinson's disease; immune disorder; haematopoietic disorder;
 KW haemophilia; hypercoagulation; Crohn's disease; cancer.

XX OS Homo sapiens.

XX PN WO2002021498-A2.

XX PD 17-OCT-2002.

XX PF 03-APR-2002; 2002WO-US010780.

XX PR 03-APR-2001; 2001US-0281086P.

PR 03-APR-2001; 2001US-0281136P.

PR 05-APR-2001; 2001US-0281863P.

PR 06-APR-2001; 2001US-0281906P.

PR 10-APR-2001; 2001US-0282020P.

PR 10-APR-2001; 2001US-0282934P.

PR 12-APR-2001; 2001US-0283512P.

PR 13-APR-2001; 2001US-0283710P.

PR 17-APR-2001; 2001US-0284234P.

PR 20-APR-2001; 2001US-0285325P.

PR 20-APR-2001; 2001US-0285609P.

PR 23-APR-2001; 2001US-0285748P.

PR 23-APR-2001; 2001US-0285890P.

PR 24-APR-2001; 2001US-0286068P.

PR 25-APR-2001; 2001US-0286292P.

PR 27-APR-2001; 2001US-0287213P.

PR 02-MAY-2001; 2001US-0288257P.

PR 30-MAY-2001; 2001US-0294164P.

PR 30-MAY-2001; 2001US-0294484P.

PR 18-JUN-2001; 2001US-0298952P.

PR 19-JUN-2001; 2001US-0299237P.

PR 19-JUN-2001; 2001US-0299276P.

PR 12-SEP-2001; 2001US-0318750P.

PR 25-SEP-2001; 2001US-0324800P.
 PR 25-SEP-2001; 2001US-0324802P.
 PR 27-SEP-2001; 2001US-0325684P.
 PR 17-OCT-2001; 2001US-0330143P.
 PR 14-NOV-2001; 2001US-0332131P.
 PR 14-NOV-2001; 2001US-0332240P.
 PR 14-NOV-2001; 2001US-0332779P.
 PR 21-NOV-2001; 2001US-0332115P.
 PR 04-DEC-2001; 2001US-0337621P.
 PR 03-JAN-2002; 2002US-0345783P.
 PR 16-JAN-2002; 2002US-0350251P.
 PR 02-APR-2002; 2002US-00114270.
 XX (CURA-) CURAGEN CORP.
 XX Guo X, Kekuda R, Miller CE, Malyankar UM, Spytek KA;
 PI Patturajan M, Liu X, Gusev VY, Li L, Vernet CM, Zerhusen BD;
 PI Gorman L, Shenoy SG, Pena CE, Smithson G, Burgess CE, Gerlach V;
 PI Padigaru M, Shimkets RA, Gangolli EA, Taupier RJ, Casman SU, Ji W;
 PI Anderson DW, Leite MW, Rastelli L, Edinger SR, Stone DJ;
 PI MacDougall JR, Rothenberg ME, Mazur A, Millet I, Peyman JA;
 PI Ellerman K;
 XX WPI; 2003-046858/04.
 DR P-PSDB; ABUS4635.
 XX New isolated NOVX polypeptide useful for treating atherosclerosis,
 PT metabolic disorders, diabetes, obesity, infectious disease, anorexia,
 PT neurodegenerative disorders, Alzheimer's disease and cancer.
 XX Claim 17; Page 291-292; 666pp; English.
 XX The invention relates to human polypeptides, termed NOVX, and the
 CC polynucleotides encoding them. The polypeptides and polynucleotides are
 CC useful for diagnosing disease, and screening for potential therapeutic
 CC agents. The sequences are useful for treating metabolic disorders,
 CC cardiomyopathy, diabetes, hypertension, congenital heart defects, aortic
 CC stenosis, atrial septal defect (ASD), atrioventricular canal defect,
 CC ductus arteriosus, pulmonary stenosis, subaortic stenosis, ventricular
 CC septal defect (VSD), valve diseases, tuberosus sclerosis, scleroderma,
 CC atherosclerosis, obesity, infectious disease, anorexia, neurodegenerative
 CC disorders, Alzheimer's disease, Parkinson's disease, immune disorders,
 CC haematopoietic disorders, haemophilia, hypercoagulation, Crohn's disease
 CC and cancer. Sequences ABX72170-ABX72275 represent human NOVX
 CC polynucleotides of the invention

XX SQ Sequence 1587 BP; 513 A; 313 C; 347 G; 414 T; 0 U; 0 Other;

Query Match 50.6%; Score 1199.6; DB 8; Length 1587;

Best Local Similarity 99.4%; Pred. No. 8.6e-307;

Matches 1215; Conservative 0; Mismatches 4; Indels 3; Gaps 1;

QY 1 CATTGGAGACCATGATAGTACGATGTGATTAGCCATCGGCAAGTGCCTTCGGGA 60

DB 88 CATTGGAGACCATGATAGTACGATGTGATTAGCCATCGGCAAGTGCCTTCGGGA 147

QY 61 AAGCATCTTAGCTAAAGGGAATCAGATAGCAAGCACTGTGTATAAAGAGATCAATT 120

DB 148 AAGCATCTTAGCTAAAGGGAATCAGATAGCAAGCACTGTGTATAAAGAGATCAATT 207

QY 121 TTGAAAGATGCCCATACAAGAAAAGAGCTTCAAGAAAAGATGATTTCTTCTGGAAA 180

DB 208 TTGAAAGATGCCCATACAAGAAAAGAGCTTCAAGAAAAGATGATTTCTTCTGGAAA 267

QY 181 AGATGAACATCCCAACATTTAGCTTCTTCAATTTCAAGAGATGGCAGGCTGT 240

DB 268 AGATGAACATCCCAACATTTAGCTTCTTCAATTTCAAGAGATGGCAGGCTGT 327

QY 241 TTATTGTAAATGGAATATTGTGATGAGGGGATCTCATGAAAGGATCAATAGACACGG 300

DB 328 TTATTGTAAATGGAATATTGTGATGAGGGGATCTCATGAAAGGATCAATAGACACGG 387

QY 301 GTGTGTATTAGTGAAGATCAGATCCCTCGGTGGTTGTACAGATTTCTCTAGGACTAA 360

Db 388 GTGTGTTATTAGTGAAGATCAGATCCTCGGTGGTTTGTACAGATTCTCTAGCACTAA 447
Qy 361 AACATATTTCATGACAGGAAGATATTACACAGGACATAAAAGCTCAGAACATTTTCTTA 420
Db 448 AACATATTTCATGACAGGAAGATATTACACAGGACATAAAAGCTCAGAACATTTTCTTA 507
Qy 421 GCAAGAACGGAATGGTGGCAAGCTTGGGACATTTGGTATAGCAAGAGTCTGAATAATT 480
Db 508 GCAAGAACGGAATGGTGGCAAGCTTGGGACATTTGGTATAGCAAGAGTCTGAATAATT 567
Qy 481 CCATGGAATCTGCTCGAACTTGATTGGAACACCTTACTACCTGTCCTCCAGAGATCTGTC 540
Db 568 CCATGGAATCTGCTCGAACTTGATTGGAACACCTTACTACCTGTCCTCCAGAGATCTGTC 627
Qy 541 AGAATAAACCCCTACAACTAAATAAAGGATATTGGTCTCTTGGCTGTGCTTATATAGGC 600
Db 628 AGAATAAACCCCTACAACTAAATAAAGGATATTGGTCTCTTGGCTGTGCTTATATAGGC 687
Qy 601 TCTGCACTATTAAACATCTTTGAGGTAACACTTACAGACAGCTGTTCTGAGATTT 660
Db 688 TCTGCACTATTAAACATCTTTGAGGTAACACTTACAGACAGCTGTTCTGAGATTT 747
Qy 661 GTCAAGCACATTTTGCCCAATATCTCCGGGTTTTCTCGTGTGAGTCCATTCCTTGATAT 720
Db 748 GTCAAGCACATTTTGCCCAATATCTCCGGGTTTTCTCGTGTGAGTCCATTCCTTGATAT 807
Qy 721 CTCAGCTCTTCAAGTATCTCTCGAGACCGACCATTCATTAATTTCCATTTGAAAGGC 780
Db 808 CTCAGCTCTTCAAGTATCTCTCGAGACCGACCATTCATTAATTTCCATTTGAAAGGC 867
Qy 781 CTTTTTAGAGATCTTATTTCCAAATATTTCGACTCTCGAGTCA--TTCAAGGAAGAT 837
Db 868 CTTTTTAGAGATCTTATTTCCAAATATTTCGACTCTCGAGTCA--TTCAAGGAAGAT 927
Qy 838 TCAGTCACATGCTTATATGACAGACGAGCGCCAGCTTCTCGAATGCTGGGAAGTGG 897
Db 928 TCAGTCACATGCTTATATGACAGACGAGCGCCAGCTTCTCGAATGCTGGGAAGTGG 987
Qy 898 TCAGGAAGTAAATACAAAAGTGAATTCGGGGAAGTCCCAACCAATCAAGATCAAGGA 957
Db 988 TCAGGAAGTAAATACAAAAGTGAATTCGGGGAAGTCCCAACCAATCAAGATCAAGGA 1047
Qy 958 TATCTGTGCAATTAAGGAATGCTTATATGATAGAAATGAATGGAGACCAACAGCTG 1017
Db 1048 TATCTGTGCAATTAAGGAATGCTTATATGATAGAAATGAATGGAGACCAACAGCTG 1107
Qy 1018 GAGCCAGAGCGCAGATCTATAAAATGATAGAAAGCCCAAAATGCTGCTGTCTGTG 1077
Db 1108 GAGCCAGAGCGCAGATCTATAAAATGATAGAAAGCCCAAAATGCTGCTGTCTGTG 1167
Qy 1078 GACATTATGATTATTATTATGCTCAACTGATATGCTGAGGAGAGAGCCCAACCA 1137
Db 1168 GACATTATGATTATTATTATGCTCAACTGATATGCTGAGGAGAGAGCCCAACCA 1227
Qy 1138 GTTATCACCTATCTTCAAGAAATATCTGGAGTTACGGTCAAGAAACAGGC 1197
Db 1228 GTTATCACCTATCTTCAAGAAATATCTGGAGTTACGGTCAAGAAACAGGC 1287
Qy 1198 ATGGTCCATCCCAAGTCAATG 1219
Db 1288 ATGGTCCATCCCAAGTCAATG 1309

RESULT 10

ABX72262

ID ABX72262 standard; cDNA; 1453 BP.

XX AC ABX72262;

XX DT 03-JUN-2003 (first entry)

XX DE Human NOVX polynucleotide #93.

XX Human; NOVX; gene; ss; metabolic disorder; cardiomyopathy; diabetes; ASD;
KW hypertension; congenital heart defect; aortic stenosis; valve disease;
KW atrial septal defect; atrioventricular canal defect; ductus arteriosus;
KW pulmonary stenosis; subaortic stenosis; ventricular septal defect; VSD;
KW tuberculous sclerosis; scleroderma; atherosclerosis; infectious disease;
KW obesity; anorexia; neurodegenerative disorder; Alzheimer's disease;
KW Parkinson's disease; immune disorder; haematopoietic disorder;
KW haemophilia; hypercoagulation; Crohn's disease; cancer.
XX Homo sapiens.
XX WO200281498-A2.
XX 17-OCT-2002.
XX 03-APR-2002; 2002WO-US010780.
XX 03-APR-2001; 2001US-0281086P.
XX 03-APR-2001; 2001US-0281136P.
XX 05-APR-2001; 2001US-0281883P.
XX 05-APR-2001; 2001US-0281906P.
XX 06-APR-2001; 2001US-0282020P.
XX 10-APR-2001; 2001US-0282930P.
XX 10-APR-2001; 2001US-0282934P.
XX 12-APR-2001; 2001US-0283512P.
XX 13-APR-2001; 2001US-0283710P.
XX 17-APR-2001; 2001US-0284234P.
XX 19-APR-2001; 2001US-0285325P.
XX 20-APR-2001; 2001US-0285381P.
XX 20-APR-2001; 2001US-0285609P.
XX 23-APR-2001; 2001US-0285748P.
XX 23-APR-2001; 2001US-0285890P.
XX 24-APR-2001; 2001US-0286086P.
XX 25-APR-2001; 2001US-0286292P.
XX 27-APR-2001; 2001US-0287213P.
XX 02-MAY-2001; 2001US-0288257P.
XX 29-MAY-2001; 2001US-0294164P.
XX 30-MAY-2001; 2001US-0294484P.
XX 18-JUN-2001; 2001US-0298952P.
XX 19-JUN-2001; 2001US-0299237P.
XX 19-JUN-2001; 2001US-0299276P.
XX 12-SEP-2001; 2001US-0318750P.
XX 25-SEP-2001; 2001US-0324800P.
XX 25-SEP-2001; 2001US-0324802P.
XX 27-SEP-2001; 2001US-0325684P.
XX 17-OCT-2001; 2001US-0330143P.
XX 14-NOV-2001; 2001US-0332131P.
XX 14-NOV-2001; 2001US-0332240P.
XX 21-NOV-2001; 2001US-0332779P.
XX 04-DEC-2001; 2001US-0332115P.
XX 03-JAN-2002; 2001US-0337621P.
XX 16-JAN-2002; 2002US-0345783P.
XX 02-APR-2002; 2002US-0350251P.
XX (CURA-) CURAGEN CORP.
XX Guo X. Kekuda R. Miller CE, Malvankar UM, Spytek KA;
PI Patturajan M, Liu X, Gusev VV, Li L, Vernet CAM, Zerhusen BD;
PI Gorman L, Shenoy SG, Pena CE, Smithson G, Burgess CE, Gerlach V;
PI Padigaru M, Shinkets RA, Gangoli EA, Taupier RJ, Casman SJ, Ji W;
PI Anderson DW, Leite MW, Rastelli L, Edinger SR, Stone DU;
PI Macdougall JR, Rothenberg ME, Mazur A, Millet I, Peyman JA;
PI Ellerman K;
XX WPI; 2003-046858/04.
XX P-FSDB; ABUS4634.
XX New isolated NOVX polypeptide useful for treating atherosclerosis,
XX metabolic disorders, diabetes, obesity, infectious disease, anorexia,
XX neurodegenerative disorders, Alzheimer's disease and cancer.

Claim 17; Page 291; 666pp; English.

The invention relates to human polypeptides, termed NOVX, and the polynucleotides encoding them. The polypeptides and polynucleotides are useful for diagnosing disease, and screening for potential therapeutic agents. The sequences are useful for treating metabolic disorders, cardiomyopathy, diabetes, hypertension, congenital heart defects, aortic stenosis, atrial septal defect (ASD), atrioventricular canal defect, ductus arteriosus, pulmonary stenosis, subaortic stenosis, ventricular septal defect (VSD), valve diseases, tuberosus sclerosis, scleroderma, atherosclerosis, obesity, infectious disease, anorexia, neurodegenerative disorders, Alzheimer's disease, Parkinson's disease, immune disorders, haematopoietic disorders, haemophilia, hypercoagulation, Crohn's disease and cancer. Sequences ABX72170-ABX72275 represent human NOVX polynucleotides of the invention.

Sequence 1453 BP; 466 A; 283 C; 309 G; 395 T; 0 U; 0 Other;

Query Match 49.7%; Score 1176.8; DB 8; Length 1453;

Best Local Similarity 99.1%: Pred. No. 9e-30f:

Best local similarity 55.1%, seq. no. 38-301,
Matches 1197; Conservative 0; Mismatches 2; Indels 9; Gaps 1;

Qy	12	ATGGATTAAGTACGATGTGATTAAAGCCNATCGGGCAAGTGCCTTCGGGAAAGCATCTTAA	71
Db	1	ATGGATAAGTACGATGTGATTAAAGCCNATCGGGCAAGTGCCTTCGGGAAAGCATCTTAA	60
Qy	72	GCTTAAAGGGAATCAGATAGCAAGCACCTGTGTCATAAAAGAGATCAATTTTGAAAAGATG	131
Db	61	GCTTAAAGGGAATCAGATAGCAAGCACCTGTGTCATAAAAGAGATCAATTTTGAAAAG---	117
Qy	132	CCCATACAAGAAAAAGAGCTTCAAAGAAAGAGTGAATCTTCTGGAAAAGATGAAACAT	191
Db	118	-----CAAGAAAAAGAGCTTCAAAGAAAGAGTGAATCTTCTGGAAAAGATGAAACAT	171
Qy	192	CCCAACATTGTAGCTTCTTCAAATTCATTTCAGAGAGAAATGCGAGGCTGTTATTTGTAATG	251
Db	172	CCCAACATTGTAGCTTCTTCAAATTCATTTCAGAGAAATGCGAGGCTGTTATTTGTAATG	231
Qy	252	GAATATTGTGATGGAGGGGATCTCATGAAAGAGATCAATATAGACAACCGGGGTGTATTATT	311
Db	232	GAATATTGTGATGGAGGGGATCTCATGAAAGAGATCAATATAGACAACCGGGGTGTATTATT	291
Qy	312	AGTGAAGATCAGATCCTCGGTGGTTGTACAGATTTCTCTAGAGCTTAAAAACATATTTCAT	371
Db	292	AGTGAAGATCAGATCCTCGGTGGTTGTACAGATTTCTCTAGAGCTTAAAAACATATTTCAT	351
Qy	372	GCACGAGAAGATATTACAAGGACATAAAAGCTCAGAACATTTTTTCTTAGCAAGAACGGA	431
Db	352	GCACGAGAAGATATTACAAGGACATAAAAGCTCAGAACATTTTTTCTTAGCAAGAACGGA	411
Qy	432	ATGGTGGCAAGACTTTGGGACCTTTGGTATAGCAAGAGTCTCTGAATTAATTCATCGAACATT	491
Db	412	ATGGTGGCAAGACTTTGGGACCTTTGGTATAGCAAGAGTCTCTGAATTAATTCATCGAACATT	471
Qy	492	GCTCGAACTTGTATTGGAACACTTACTACTCTGCCCAGAGATCTGTACAGATAAACCC	551
Db	472	GCTCGAACTTGTATTGGAACACTTACTACTCTGCCCAGAGATCTGTACAGATAAACCC	531
Qy	552	TACACAATAAACCAGATATTGGTGCTCTGGCTGTGCTTATATAGAGCTCTGCACACTT	611
Db	532	TACACAATAAACCAGATATTGGTGCTCTGGCTGTGCTTATATAGAGCTCTGCACACTT	591
Qy	612	AAACATCTCTTTGAGGGTAACAACTTACAGACGTGGTTCTGAAGATTTGTCTAAGCACAT	671
Db	592	AAACATCTCTTTGAGGGTAACAACTTACAGACGTGGTTCTGAAGATTTGTCTAAGCACAT	651
Qy	672	TTTGCCCCCAATATCTCCGGGGTTTTCTCGTACGTCCATTCCTTTGATATCTACAGCTTTT	731
Db	652	TTTGCCCCCAATATCTCCGGGGTTTTCTCGTACGTCCATTCCTTTGATATCTACAGCTTTT	711
Qy	732	CAAGTATCTCTCGAGACCGACCATCCATAAATTCATTTTGAAAAGGCCCTTTTAGAG	791
Db	712	CAAGTATCTCTCGAGACCGACCATCCATAAATTCATTTTGAAAAGGCCCTTTTAGAG	771

```
XX WPI: 2003-201429/19.
DR P-PSDB; ABP96073.
XX
PT New protein kinase genes and polypeptides, useful for diagnosing diseases
PT associated with a protein kinase, or in gene therapy for treating e.g.
PT Parkinson's disease, migraine, myocardial infarction, allograft rejection
PT or cancers.
PT
XX
XX Claim 1; Page 76; 258pp; English.
PS
XX
XX ABZ77126 to ABZ77165 encode the human protein kinases given in ABP96048
XX to ABP96087. The protein kinases have antiasthmatic, antiinflammatory,
XX antidiabetic, antiparkinsonian, antimigraine, cardiant, cytostatic,
XX immunosuppressive and vulnary activities, and can be used in gene
XX therapy. A protein kinase therapeutic agent from the present invention,
XX particularly a protein kinase gene agonist or antagonist, can be used for
XX treating a disease or condition associated with a protein kinase in an
XX individual. These diseases include chronic obstructive pulmonary diseases
XX (COPD), asthma, non-insulin dependent diabetes, Parkinson's disease,
XX migraine, myocardial infarction, inflammatory bowel disease, autoimmune
XX disorders (e.g. allograft rejection or graft vs. host disease), cancers
XX (e.g. leukemias) or wound granulation
XX
XX Sequence 1632 BP; 459 A; 354 C; 415 G; 404 T; 0 U; 0 Other;
XX
XX Query Match 25.4%; Score 602; DB 10; Length 1632;
XX Best Local Similarity 98.4%; Pred. No. 1.6e-148;
XX Matches 608; Conservative 0; Mismatches 10; Indels 0; Gaps 0;
XX
QY 215 TTCATTTCAAGAGATGGCAGCGCTGTTATTTGTAATGGAATATTGTGATGAGGGGATCT 274
DB ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 275 CATGAAAGGATCAATAGACACGGGGTGTGTTATTTAGTGAATCATGATCCCTCGGTTG 334
DB ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 261 CATGAAAGGATCAATAGACACGGGGTGTGTTATTTAGTGAATCATGATCCCTCGGTTG 320
QY 335 GTTGTACAGATTTCTTAGGACTAAACATATTATCATGACGGAAGATATTACACAGGGA 394
DB ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 321 GTTGTACAGATTTCTTAGGACTAAACATATTATTCATGACGGAAGATATTACACAGGGA 380
QY 395 CATAAAGCTCAGACATTTTCTTAGCAAGACGGAATGGTGGCAAGCTTGGGGACTT 454
DB ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 381 CATAAAGCTCAGACATTTTCTTAGCAAGACGGAATGGTGGCAAGCTTGGGGACTT 440
QY 455 TGGTATAGCAAGAGCTCTGAATAATTCCATGGAATTTGCTGAACTTGTATTGGAACACC 514
DB ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 441 TGGTATAGCAAGAGCTCTGAATAATTCCATGGAATTTGCTGAACTTGTATTGGAACACC 500
QY 515 TTACTACCTGTCCAGAGATCTGTGAGATTAACCCCTACACATATAAAGCGATATTG 574
DB ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 501 TTACTACCTGTCCAGAGATCTGTGAGATTAACCCCTACACATATAAAGCGATATTG 560
QY 575 GTCTCTGGCTGTCTTATATGAGCTCTGCACATCTTAAACATCCTTTGAGGGTAACAA 634
DB ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 561 GTCTCTGGCTGTCTTATATGAGCTCTGCACATCTTAAACATCCTTTGAGGGTAACAA 620
QY 635 CTTACAGAGCTGTCTGAAGATTGTCAAGCAATTTGCCCCCAATATCTCCGGGTT 694
DB ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 621 CTTACAGAGCTGTCTGAAGATTGTCAAGCAATTTGCCCCCAATATCTCCGGGTT 680
QY 695 TTCTCGTAGCTCCATTCCTTCATATCTAGCTCTTTCAAGTATCTCTCCGAGCCGACC 754
DB ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 681 TTCTCGTAGCTCCATTCCTTCATATCTAGCTCTTTCAAGTATCTCTCCGAGCCGACC 740
QY 755 ATCCATAAATCCATTTTGAAGGCCCTTTTAGAGATCTTTATCCCAATATTTCAC 814
DB ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 741 ATCCATAAATCCATTTTGAAGGCCCTTTTAGAGATCTTTATCCCAATATTTCAC 800
QY 815 TCCTGAGTCAATCAGGA 832
DB ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 801 TCCTGAGTCAATTTTGA 818
```

RESULT 12

ADP44502
ID ADP44502 standard; cDNA; 1434 BP.

XX AC ADP44502;

XX DT 12-FEB-2004 (first entry)

XX DE Mouse kinase protein encoding cDNA SEQ ID NO:20.

XX KW cytosolic; nootropic; neuroprotective; antidiabetic; screening;
XX KW regulation; drug development; protein-associated disease; cancer;
XX KW dementia; diabetes; kinase; enzyme; mouse; gene; ss.

XX OS Mus musculus.

XX FN WO2003084992-A1.

XX PD 16-OCT-2003.

XX PF 04-APR-2003; 2003WO-JP004330.

XX PR 05-APR-2002; 2002JP-00103396.

XX PR 23-APR-2002; 2002JP-00120904.

XX PR 02-MAY-2002; 2002JP-00130601.

XX PR 04-DEC-2002; 2002JP-00352520.

XX PA (RIKE) RIKEN KK.

XX PA (DNAF-) DNAFORM KK.

XX PA (MITU) MITSUBISHI CHEM CORP.

XX PI Hayashizaki Y, Kamiya M, Kubodera H, Watanabe W;

XX WPI: 2003-833568/77.

XX P-PSDB; ADP44528.

Proteins and encoded DNAs with kinase activity, useful in screening substances for regulating such activity and in developing drugs for the protein-associated diseases e.g. cancer, dementia and diabetes.

Claim 4; SEQ ID NO 20; 342pp; Japanese.

The present invention describes a protein: (a) containing any of the amino acid sequences of ADP44509 to ADP44534 or ADP44544; or (b) based on any of the sequences in (a) but with some amino acids deleted, substituted and/or added and having kinase activity. Also described: (1) a DNA encoding any of the proteins; (2) a full-length cDNA encoding the protein; (3) a DNA which is: (a) a DNA containing any of the base sequences in ADP44483 to ADP44508 or ADP44543; (b) a DNA derived from any of the sequences in (a) but with some bases deleted, substituted and/or added and encoding a protein with kinase activity; or (c) a DNA hybridisable with any of the sequences in (a) or their complementary strands under stringent conditions and encoding a protein with kinase activity; (4) a recombinant vector containing the DNA; (5) a cell transfected with the DNA or recombinant vector, or an individual produced from the cell; (6) recombinant proteins produced by such cells; (7) an oligonucleotide containing 5-10 consecutive bases in any of the base sequences, its sense oligonucleotide, an antisense oligonucleotide with a complementary strand of such sense oligonucleotide, or an oligonucleotide derivative of the (anti-)sense oligonucleotide; (8) an antibody specifically binding to the protein, or its partial fragment; (9) a method for screening substances for regulating activity of the protein by contacting a test substance with such protein before measuring changes in the protein activity due to the test substance; (10) a method for screening substances regulating expression of the DNA by contacting a test substance with cells transfected with the gene and detecting changes in expression level of the DNA in such cells; (11) recordable media for reading in a computer with information on the amino acid sequences of the proteins, and/or base sequences of the DNAs scored; and (12) a support for binding with any of the proteins and/or DNAs. The proteins and their encoded DNAs have cytostatic, nootropic, neuroprotective and antidiabetic

CC activities. They can be used in screening substances for regulating such
CC activity and in developing drugs for the protein-associated diseases e.g.
CC cancer, dementia and diabetes. The present sequence is used in the
XX exemplification of the present invention.

QQ Sequence 1434 BP; 392 A; 338 C; 336 G; 368 T; 0 U; 0 Other;

Query Match 25.2%; Score 597.4; DB 10; Length 1434;
Best Local Similarity 80.3%; Pred. No. 2.4e-147;
Matches 730; Conservative 0; Mismatches 166; Indels 13; Gaps 2;

QY 10 CCATGGATAGTAGTGTGATTAAGCCATCGGCAAGTGTGCTTCGGGAAAGCATACT 69
DB 102 CCATGGATAACTTTCACTGATTAAATCATTTGGGAAGCACCTTTGGGAAGTGTACT 161
QY 70 TAGCTAAAGGAATCAGATAGCAAGCACTGTGTCTATAAAGAGATCAATTTTGAAGA 129
DB 162 TGGCTAAAGATAAATCAGAAAGCAGTCACTGTGTCTATAAAGAAATCAGTTTCA 220
QY 130 TGCCCATACAGAAAGAAAGAGCTTCAAGAAAGAGTATTCTTCTGAAAGATGAAAC 189
DB 221 -----GAAAGAGAGCCCTCAAGAACAGTGTCTTCTGCTAGATGGAGC 269
QY 190 ATCCCAATGTAGCTTCTTCAATTCATTCAAGAGATGCGAGCTGTTTATTGTAA 249
DB 270 ATCCCAATGTAGCTTCTTCAATTCATTCAAGAGATGCGAGCTGTTTATTGTAA 329
QY 250 TGGAAATTTGTGAGGAGGATCTCATGAAAGGATCAATAGACACGGGGTGTGTAT 309
DB 330 TGGAAATTTGTGAGGAGGATCTCATGAGAGATCCAGAGCGCGGGAGTGTATGT 389
QY 310 TTAGTGAAGATCAGATCCTCGTGTGTTGTACAGATTTCTTAGGACTAAACATATTC 369
DB 390 TCAGCAAGACCATCTGTGTGTTGTACAGATTTCTTAGGACTGAAGCATATTC 449
QY 370 ATGACAGAGATATTACAGGAGACATAAAGCTCAGACATTTTCTTAGCAAGACG 429
DB 450 ATGACAGAGATTTTACAGGAGACATAAATCTCAGAAATATTTTCTTAGCAAGATG 509
QY 430 GAATGCTGGCAAGCTTGGGACATTTGGTATAGCAAGAGTCTCAATATTCATGGAAC 489
DB 510 GAATGCTGGCAAGCTTGGGACATTTGGGACATTTGGGACATTTGGGACATTTGGG 569
QY 490 TTGCTGAACTTTGGAACACCTTACTACCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 549
DB 570 TTGCTGAACTTTGGAACACCTTACTACCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 629
QY 550 CCTACAACTTTGGAACACCTTACTACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 609
DB 630 CATACAACTTTGGAACACCTTACTACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 689
QY 610 TTAACATCTTTTGGAGGTAACAACTTACAGAGCTGGTTCGAGATTTGTCAGAGCAC 669
DB 690 TCAAGCATCTTTTGGAGGTAACAACTTACAGAGCTGGTTCGAGATTTGTCAGAGCAC 749
QY 670 ATTTTGGCCCAATCTCCGGGGTTTCTCGTAGCTCCATCTCTGATATCTCAGTCT 729
DB 750 GTGTTGCTCCATATCAACCACTTCTCTGACCTACAGTCTTGTATACCTCAGCTCT 809
QY 730 TTCAAGTATCTCTCGAGACCGACCATCCATAATTTCAATTTTGAAGAGCCCTTTTAT 789
DB 810 TCAGAGTGTCTCTCAGACCGGACATCTCTAGTCTCTTTTGAAGAGCCCTTTTAT 869
QY 790 AGAATCTTATCCCAATATTTGACTCTCTGAGTCTCTGAGTCTCTGAGTCTCTGAGT 848
DB 870 AATCTCTATCTCCGAGTCTCTGAGTCTCTGAGTCTCTGAGTCTCTGAGTCTCTGAG 929
QY 849 CTTATATGACAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 908
DB 930 CTCACATGGAGAAATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 989
QY 909 AANATCAA 917

DB 990 ATCATGTAA 998

RESULT 13

ADP44489

ID ADF44489 standard; cDNA; 2407 BP.

XX

AC ADF44489;

XX

DT 12-FEB-2004 (first entry)

XX

Mouse kinase protein encoding cDNA SEQ ID NO:7.

XX

cytostatic; nootropic; neuroprotective; antidiabetic; screening;
regulation; drug development; protein-associated disease; cancer;
dementia; diabetes; kinase; enzyme; mouse; gene; ss.

XX

Mus musculus.

OS

WO2003084992-A1.

XX

16-OCT-2003.

XX

04-APR-2003; 2003WO-JP004330.

XX

05-APR-2002; 2002JP-00103396.

XX

23-APR-2002; 2002JP-00120904.

PR

02-MAY-2002; 2002JP-00130601.

PR

04-DEC-2002; 2002JP-00352520.

XX

(RIKE) RIKEN KK.

PA

(DNAP-) DNAFORM KK.

PA

(MITU) MITSUBISHI CHEM CORP.

XX

Hayashizaki Y, Kamiya M, Kubodera H, Watanabe W;

XX

WPI; 2003-833568/77.

DR

P-PSDB; ADF44515.

XX

Proteins and encoded DNAs with kinase activity, useful in screening

PT

substances for regulating such activity and in developing drugs for the

PT

protein-associated diseases e.g. cancer, dementia and diabetes.

XX

Claim 4; SEQ ID NO 7; 342bp; Japanese.

PS

XX

XX

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC encoded DNAs have cytostatic, nootropic, neuroprotective and antidiabetic
CC activities. They can be used in screening substances for regulating such
CC activity and in developing drugs for the protein-associated diseases e.g.
CC cancer, dementia and diabetes. The present sequence is used in the
CC exemplification of the present invention.

XX
SQ Sequence 2407 BP; 718 A; 528 C; 604 G; 557 T; 0 U; 0 Other;

Query Match 25.2%; Score 597.4; DB 10; Length 2407;
Best Local Similarity 81.1%; Pred. No. 3.1e-147;
Matches 724; Conservative 0; Mismatches 156; Indels 13; Gaps 2;
QY 10 CCATGTAAGTACGATGATTAAGGCCATCGGCAAGTCCCTTCGGGAAGCATACT 69
DB 90 CCATGTAAGTACGATGATTAAGGCCATCGGCAAGTCCCTTCGGGAAGTACT 149
QY 70 TAGCTAAGGGAAATCAGATAGCAGACTGTGTATAAAGAGATCAATTTTGAAGA 129
DB 150 TGGCTAAAGATAAATCAGAAGCAGTCACTGTGTATAAAGAAATCAGTTTGACAAAG- 208
QY 130 TGCCCATACAAGAAAAGAGCTTCAAGAAAAGAGTATCTTCTGGAAGAGATGAAC 189
DB 209 -----GAAAAGAGGCTCAAGACAGAGTATCTTCTGGCTAGGATGGAGC 257
QY 190 ATCCCAACATGTAGCTTCTTCAATTCATTTCAAGAGAAATGGCAGGCTTTATTGTAA 249
DB 258 ATCCCAATATCGTAACCTTCTCAGCTCGTTTCAAGAGAAACGCGAGGCTTTATTGTAA 317
QY 250 TGGATATTGTATGGAGGGATCTCATGAAGAGATCAATAGACAACGGGGTGTATT 309
DB 318 TGAATACTGTATGAGGGGATCTCATGAGAGATCCAGGAGCGGGGAGTGATGT 377
QY 310 TTAGTGAAGATCAGATCTCTCGGTGGTTTGTACAGATTTCTTAGGACTAAACATATTC 369
DB 378 TCAGCGAAGACAGATCTCTGTGTGGTTGTACAGATTTCTTAGGACTGAAGCATATTC 437
QY 370 ATGACAGGAAGATATACACAGGACATAAAGAGCTCAGAACATTTTCTTACGACAGACG 429
DB 438 ATGACAGGAAGATTTTACACAGGACATAAATCTCAGAATATTTTCTTACGACAGATG 497
QY 430 GAATGTGCAAGCTTGGGAGCTTTGGTATAGCAAGAGTCTGATATATTCATGGAGAC 489
DB 498 GAATGTGTCAGCTTGGGAGCTTTGGACAGAGACACTGAATGATCTCATGGAGAC 557
QY 490 TTGCTCGAATGTATTTGAACACCTTACTACTGTCCTCCAGAGATCTGTGAGATAAATAC 549
DB 558 TTGCTCAAAACATGTCTGGGACACCTTACTACTGTCCTCCAGAGATCTGTGAGACAGGC 617
QY 550 CCTACACAAATAAAGCGGATATTTGGTCTCTTGGCTGTCTTATATGAGCTCTGCACAC 609
DB 618 CATACAAATAAAGCGGACATCTGGTCTCTTGGCTGTCTTATGAGCTCTGCACAC 677
QY 610 TTAACATCTTTTGAAGGTAACAACTTACAGCAGCTGGTCTGGAAGATTTTGTCAAGCAC 669
DB 678 TCAAGCATCTTTTGAAGCAACAACCTCCACATCTGTTCTGAAGATTTGTCAAGGAC 737
QY 670 AITTTGCCCATATCTCGGGGTTTCTCGTGAGCTCCATTCCTTGTATCTCAGCTCT 729
DB 738 GTGTGCTCCCATATCACCCCACTTCTCTCGTGAOCTACAGTCCCTTGTATCTCAGCTCT 797
QY 730 TTCAAGTATCTCTCGAGACCGACCATCCATAAATTCATTTTGAAGGCGCTTTTGTAG 789
DB 798 TCAGAGTGTCTCTCAGACCGGCGCATCGTTACGTCCTTTTGAAGACCGCTTTTGTAG 857
QY 790 AGAATCTTATTTCCAAATATTTGACTCTCTGAGGTC-AITTCAGAGAAATTCAGTCAATG 848
DB 858 AAATCTCTATTTGCCCATCTTGTATCTCTGAGGTCGTGTTTCAAGAGAAATTCAGTCCACG 917
QY 849-CTTATATGAGACGAGCGCGGAGCTTCTCCAGATGCTGGGAAGTGTCTCA 901
DB 918 CTCACATGAGAGACATGGCCATTGGCCCCACAGCTTGTGGAGAGTTAGTCCA 970

RESULT 14

ABX13160
XX ID ABX13160 standard; cDNA; 4683 BP.
XX AC ABX13160;
XX DT 10-MAY-2003 (first entry)
XX Human cDNA for kinase and phosphatase, KPP-14, INCYTE No. 8043337CBI.
DE Human; ss; gene; kinase and phosphatase; KPP; cancer; cirrhosis;
XX cell proliferative disorder; arteriosclerosis; atherosclerosis;
XX hepatitis; paroxysmal nocturnal haemoglobinuria; polycythaemia vera;
KW psoriasis; primary thrombocytopaenia; developmental disorder;
KW renal tubular acidosis; anaemia; mental retardation; AIDS; epilepsy;
KW neurological disorder; Alzheimer's disease; Parkinson's disease;
KW autoimmune disorder; inflammatory disorder; allergy; asthma;
KW acquired immunodeficiency syndrome; autoimmune thyroiditis;
KW contact dermatitis; Crohn's disease; diabetes mellitus;
KW glomerulonephritis; Goodpasture's syndrome; gout; Graves' disease;
KW Hashimoto's thyroiditis; irritable bowel syndrome; multiple sclerosis;
KW osteoarthritis; osteoporosis; pancreatitis; Reiter's syndrome;
KW rheumatoid arthritis; Sjogren's syndrome; uveitis; infection.
XX Homo sapiens.
OS
XX WO2003012065-A2.
XX 13-FEB-2003.
XX 01-AUG-2002; 2002WO-US024521.
XX 02-AUG-2001; 2001US-0309627P.
PR 07-AUG-2001; 2001US-0310933P.
PR 09-AUG-2001; 2001US-0311323P.
PR 07-SEP-2001; 2001US-0317820P.
PR 14-SEP-2001; 2001US-0322264P.
PR 28-SEP-2001; 2001US-0326098P.
PR 19-DEC-2001; 2001US-0343007P.
PR 15-MAR-2002; 2002US-036494P.
PR 24-APR-2002; 2002US-0375539P.
XX (INCY-) INCYTE GENOMICS INC.
XX Baughn MR, Yue H, Walia NK, He A, Au-Young JK, Lee SY;
PI Gietzen KJ, Lal PG, Elliott VS, Ison CH, Yang J, Lee EA, Li JX;
PI Emerling BM, Richardson TW, Warren BA, Hafalia AJA, Marquis JP;
XX WPI: 2003-239519/23.
DR P-PSDB; ABU08113.
XX New human kinases and phosphatases and polynucleotides, useful for
PT diagnosing, treating or preventing autoimmune or inflammatory disorders
PT (e.g. AIDS, allergy or anemia), multiple sclerosis, osteoarthritis,
PT cancer or hepatitis.
XX Claim 5; Page 194-196; 199pp; English.
PS The invention relates to an isolated polypeptide, which is a human kinase
XX and phosphatase, KPP (KPP-1 to KPP-18). Also included are the encoding
CC polynucleotides KPP NA, recombinant polynucleotide comprising a promoter
CC sequence operably linked to KPP NA, a cell transformed with the
CC recombinant polynucleotide, a transgenic organism comprising the
CC recombinant polynucleotide, an anti-KPP antibody, screening for
CC ant/agonists of KPP, screening for compounds which bind to or alter the
CC activity or expression of KPP, microarray where at least one element is
CC KPP NA, generating an expression profile of a sample containing
CC polynucleotides and an array comprising different nucleotide molecules
CC affixed in distinct physical locations on a solid substrate (where at
CC least one of the nucleotide molecules comprises a first oligonucleotide
CC or polynucleotide sequence specifically hybridisable with at least 30
CC contiguous nucleotides of a target KPP NA). The kinases and phosphatases
CC (KPP) polypeptides, polynucleotides, agonists and antagonists are useful

CC for diagnosing, treating or preventing disorders associated with aberrant
 CC expression of KPP, particularly cell proliferative disorders (e.g.
 CC arteriosclerosis, atherosclerosis, cirrhosis, hepatitis, paroxysmal
 CC nocturnal haemoglobinuria, polycythaemia vera, psoriasis, primary
 CC thrombocytopenia or cancer), developmental disorders (renal tubular
 CC acidosis, anaemia or mental retardation), neurological disorders (e.g.
 CC Alzheimer's disease, Parkinson's disease or epilepsy), autoimmune/
 CC inflammatory disorders (e.g. AIDS, acquired immunodeficiency syndrome,
 CC allergies, asthma, autoimmune thyroiditis, contact dermatitis, Crohn's
 CC disease, diabetes mellitus, glomerulonephritis, Goodpasture's syndrome,
 CC gout, Graves' disease, Hashimoto's thyroiditis, irritable bowel syndrome,
 CC multiple sclerosis, osteoarthritis, osteoporosis, pancreatitis, Reiter's
 CC syndrome, rheumatoid arthritis, Sjogren's syndrome, uveitis), or viral,
 CC bacterial, fungal, parasitic, protozoan or helminthic infections. They
 CC are also useful in the assessment of the effects of exogenous compounds
 CC on the expression of nucleic acid and aa sequences of proteins associated
 CC with KPP. The polynucleotides encoding KPP are useful for creating
 CC transgenic animals to model human disease. The present sequence encodes a
 CC KPP protein of the invention

XX Sequence 4683 BP; 1595 A; 827 C; 1049 G; 1212 T; 0 U; 0 Other;

Query Match 15.6%; Score 369.4; DB 8; Length 4683;
 Best Local Similarity 66.0%; Pred. No. 1e-86;
 Matches 551; Conservative 0; Mismatches 281; Indels 3; Gaps 1;

QY 9 ACCATGATTAAGTACGATGTGATTAAAGCCATCGGCAAGTGCCTTCGGGAAAGCATAC 68
 DB |||||
 QY 617 ATCATGAGAGATGTTAGACTACAGAGATGGAGAGGTTCAATTGGAAGCCATT 676
 DB |||||
 QY 69 TTACTTAAGGGAATCAGATAGCAACACCTGTGCTATAAAGAGATCAATTTTGAAG 128
 DB |||||
 QY 677 CTGTGTTAAATCTACAGAAGATGGCAGACAGTGTGTTATCAAGGAAATTAACATCTCAAGA 736
 DB |||||
 QY 129 ATGCCCATACAGAAAGAGCTTCAAGAAAGAGTGTCTCTCGGAAAGATGAAA 188
 DB |||||
 QY 737 ATGTCCAGTAAAGAAAGAGAGAAATCAAGAGAGAGTTCAGTATTTGGCAACATGAG 796
 DB |||||
 QY 189 CATCCCAACATTTAGCGCTTCTTCAATTTCAAGAGAAATGGCAGGCTGTTTATTGTA 248
 DB |||||
 QY 797 CATCCAAATATGTCAGTATAGAGATCATTTGAAGAAATGGCTCTCTCATACATGTA 856
 DB |||||
 QY 249 ATGCAATATTGTGATGAGGGATCTCATGAAAGGATCAATAGACAAAGGGTGTGTA 308
 DB |||||
 QY 857 ATGGATTACTGTGAGGAGGGATCTGTTTAAGCGAATAAATCTCAGAAAGCGGTTTG 916
 DB |||||
 QY 309 TTTAGTGAAGATCAGATCCTCGGTTGGTTGTACAGATTTCTCTAGGACTTAAACATATT 368
 DB |||||
 QY 917 TTTCAAGAGGATCAGATTTTGGACTGGTTGTACAGATATGTTTGGCCCTGAAACATGTA 976
 DB |||||
 QY 369 CATGACAGGAGATATTACAGGGACATAAAGCTCAGACATTTTCTTACGACAGAAC 428
 DB |||||
 QY 977 CATGATAGAAAATTTCTTCATCGAGACATTAATCTCAGAACATATTTTAACTAAGAT 1036
 DB |||||
 QY 429 GGAATGTGGCAAGCTTGGGACTTTGGTATAGCAAGAGTCTGTAATATTCATCGGAA 488
 DB |||||
 QY 1037 GGAACAT---ACAATTTGAGATTTTGGATTTGCTAGAGTCTTAATAGTACTGTAGAG 1093
 DB |||||
 QY 489 CTGTCTCGAATCTGTATTGGAACACCTTACTACTGTCTCCGAGAGATCTGTGCAAGATAAA 548
 DB |||||
 QY 1094 CTGCTCGAATCTGTGATAGGAGACCCATCTACTTCTGCTCACTGAAATCTGTGAAACAAA 1153
 DB |||||
 QY 549 CCCTACAAATATAAAGCGATTTTGGTCTCTGGCTGTGCTTATATAGAGTCTCTGACA 608
 DB |||||
 QY 1154 CCTTACAAATATAAAGTGACATTTGGCTCTGGGTGTGCTTTTATAGCTGTGTACA 1213
 DB |||||
 QY 609 CTTAAACATCTTTTGAAGGTAAACAATTACAGACAGCTGGTTCTGAAGATTTGTCAAGCA 668
 DB |||||
 QY 1214 CTTAAACATCTTTTGAAGCTGGCAGTATGAAAAACCTGTACTGAAGATAATATCTGGA 1273
 DB |||||
 QY 669 CATTTTGGCCCAATATCTCCGGGGTTTTCGTGAGCTCCATTCCTTGATATCTCAGCTC 728
 DB |||||
 QY 1274 TCCTTTCCACCTGTGCTTTTGCAATATTTCTATGATCTCCGAGTGTGGTGTCTCAGTTA 1333
 DB |||||

QY 729 TTTCAAGTATCTCTCGAGACCGACCATCCATATAATTCATTTTGAAAGGCCCTTTTATA 788
 DB |||||
 QY 1334 TTTAAAGAAATCTCTAGGATAGACCATCAGTCACTCAATTCATTTGGAGAAAGGTTTATA 1393
 DB |||||
 QY 789 GAGAATCTTATTTCCCAATATTTTGACTCTCTGAGGTCATTTCCAGGAAGATTCAGTGC 843
 DB |||||
 QY 1394 GCCAACGCATGAAAGATTTCTCTCTCTCAGCTTATTCGAGAGAAATTTTGTGTC 1448
 DB |||||

RESULT 15

AAS11582

ID AAS11582 standard; cDNA; 5426 BP.

XX AAS11582;

XX 24-OCT-2001 (first entry)

XX Human cDNA encoding novel human protein, NHP #26.

XX Human; novel human protein; NHP; ss; breast cancer; prostate cancer;
 XX immunogen; antibody; gene therapy; antisense.

XX Homo sapiens.

XX WO200161016-A2.

XX 23-AUG-2001.

XX 15-FEB-2001; 2001WO-US005356.

XX 18-FEB-2000; 2000US-0183582P.

XX 22-FEB-2000; 2000US-0184014P.

XX (LEXI-) LEXICON GENETICS INC.

XX Walke DW, Hu Y, Nepomnichy B, Turner CA, Zambrowicz B;

XX WPI; 2001-502793/55.

XX Isolated nucleic acids encoding novel human proteins useful for the
 XX treatment of disease and as probes for testing and detection.

XX Disclosure; Page 68-69; 69pp; English.

XX The invention relates to novel human proteins (NHP) and the nucleic acids
 XX encoding them. The nucleic acids encode mammalian transporter proteins
 XX and are useful for the treatment (e.g. by gene therapy or antisense
 XX technology) of any of a wide variety of symptoms associated with
 XX biological disorders (e.g. breast and prostate cancer) or imbalances and
 XX as probes for the identification, selection and validation of novel
 XX molecular targets for drug discovery. The proteins may be used to raise
 XX anti-NHP antibodies. The present sequence encodes an NHP of the invention

SQ Sequence 5426 BP; 1852 A; 927 C; 1163 G; 1484 T; 0 U; 0 Other;

Query Match 15.6%; Score 369.4; DB 5; Length 5426;

Best Local Similarity 66.0%; Pred. No. 1.1e-86;

Matches 551; Conservative 0; Mismatches 281; Indels 3; Gaps 1;

QY 9 ACCATGATTAAGTACGATGTGATTAAAGCCATCGGCAAGTGCCTTCGGGAAAGCATAC 68

DB 531 ATCATGAGAGATGTTAGACTACAGAGATTTGGAGAGGTTTCAATTTGGAAGCCATT 590

QY 69 TTACTTAAGGGAATCAGATAGCAACACTGTCTCATAAAGAGATCAATTTTGAAG 128

DB 591 CTTGTTAAATCTACAGAGATGGCAGACAGTATGTTATCAAGGAAATTAACATCTCAAGA 650

QY 129 ATGCCCATACAGAAAGAGCTTCAAGAAAGAGTGTCTTCTGGAAGAGATGAAA 188

DB 651 ATGTCCAGTAAAGAAAGAGAGATCAAGGAGAGAGTTCAGTATTTGGCAACATGAG 710

QY 189 CATCCCAACATTTAGCGCTTCTTCAATTCATTTCAAGAGAAATGGCAGGCTGTTTATTGTA 248

Db	711	CATCCAAATATTGCCAGTATAGAGAAATCATTTGAAGAAAATGGCTCTCTCTACATAGTA	770
Qy	249	ATGGAAATATTGTGATGTAGGGGGATCTCATGAAAAGGATCAATAGACAACGGGGTGTGTTA	308
Db	771	ATGGATTAATCTGTAGGGAGGGATCTGTTAAACGAATAAATGCTCAGAAAGCGCTTTTG	830
Qy	309	TTTAGTGAAGATCAGATCCTCGGTGGTTGTTGTCAGATTTCTCTAGGACATAAACAATATT	368
Db	831	TTTCAAGAGATCAGATTTTGGACTGGTTGTTGTCAGATATGTTTGGCCCTGGAACATGTA	890
Qy	369	CATGACAGGAAGATATTACACAGGGACATAAAAAGCTCAGAAACATTTTTTCTTAGCAAGAAC	428
Db	891	CATGATAGAAAATCTTTCATCGAGACATTAATCTCAGAAACATATTTTAACTAAAGAT	950
Qy	429	GGAAATGTCGAAGACTTGGGACTTTGGTATAGCAAGAGTCTCTGAATAATTTCCATGGAA	488
Db	951	GGAAACAGT---ACAACCTTGAGATTTTGGAAATTCCTAGAGTTCTTAATAGTACTGTAGAG	1007
Qy	489	CTTGCTCGAACTTTGTAATTGGAAACACCTTACTACTCTGTCCCACAGAGATCTGTCAGAATAAA	548
Db	1008	CTGGCTCGAACTTGCAATAGGACCCCATACTACTTGTCACTGAAATCTGTGAAACACAAA	1067
Qy	549	CCCTACAACAATAAACCAGATATTTGGTCTCTTGGCTGTGTCCTTATATGAGCTCTGCACA	608
Db	1068	CCTTACAATAATAAAGTGCATTTGGGCTCTCGGGTGTGTCCTTTATGAGCTGTGTACA	1127
Qy	609	CTTAAACATCCTTTTGAGGGTAAACAATTACAGCAGCTGGTCTCTGAAAGATTTTGTCAAGCA	668
Db	1128	CTTAAACATGCTTTTGAAGCTGGCAGTATGAAAACCTGGTACTGGAAGATAATATCTGGA	1187
Qy	669	CATTTTCCCGCAATATCTCCGGGGTTTCTCGTGTGAGCTCCATTCCTTGTATATCTCAGTCT	728
Db	1188	TCCTTTTCCACTGTCTTTTGCATTAATTCCTATGATCTCCGCAAGTTTGGTGTCTCAGTTA	1247
Qy	729	TTTCAAGTATCTCCTCGAGACCGACCATCCATAAAATTCCAATTTTGAAGAGGCCCTTTTATA	788
Db	1248	TTTAAAGAATACTTAGGGATAGACCATCAGTCAACTCCATATTTGGAGAAGGTTTATA	1307
Qy	789	GAGAAATCTTATTCCTCAAAATATTTGACTCTCTGAGGTCAATTCAGGAAGAAATTCAGTC	843
Db	1308	GCCAAACGCAATGAAAAGTTTCTCTCTCTCCTCAGCTATTTCAGAGAAGAAATTTTGTC	1362

Search completed: November 7, 2004, 21:25:42
Job time : 1072 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 7, 2004, 21:19:48 ; Search time 219 Seconds
(without alignments)
7692.099 Million cell updates/sec

Title: US-10-730-010-1

Perfect score: 2370
Sequence: 1 cattggagaccatggataag.....ctgtcatgcgcttagcagcg 2370

Scoring table: IDENTITY NUC
Gapop 10_0 , Gapext 1.0

Searched: 824507 seqs, 355394441 residues

Total number of hits satisfying chosen parameters: 1649014

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents NA:*

1: /cgn2_6/ptodata/1/ina/5A-COMB.seq:*

2: /cgn2_6/ptodata/1/ina/5B-COMB.seq:*

3: /cgn2_6/ptodata/1/ina/6A-COMB.seq:*

4: /cgn2_6/ptodata/1/ina/6B-COMB.seq:*

5: /cgn2_6/ptodata/1/ina/PTUS-COMB.seq:*

6: /cgn2_6/ptodata/1/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2370	100.0	2370	4	US-09-873-404-1
2	2370	100.0	2370	4	US-10-243-735-1
3	369.4	15.6	5448	4	US-09-620-312D-246
4	369.4	15.6	5532	4	US-09-620-312D-245
5	355.8	15.0	4263	4	US-09-166-350-29
6	238.4	10.1	1846	3	US-09-173-581-15
7	238.4	10.1	1846	3	US-09-420-915-15
8	205.2	8.7	63588	4	US-09-873-404-3
9	205.2	8.7	63588	4	US-10-243-735-3
10	170.2	7.2	2079	4	US-09-992-481-1
11	166.8	7.0	2940	3	US-08-570-529-1
12	166.8	7.0	2940	4	US-09-544-794-1
13	166.8	7.0	3264	3	US-08-870-529-8
14	166.8	7.0	3264	4	US-09-544-794-8
15	131.6	5.6	1449	4	US-10-196-927-3
16	131.6	5.6	1938	4	US-10-196-927-1
17	131.6	5.6	2847	4	US-10-196-927-5
18	119.6	5.0	2765	4	US-09-799-451-50
19	108.6	4.6	2119	4	US-09-167-206-1
20	106.8	4.5	906	3	US-09-221-235-3
21	106.8	4.5	906	3	US-09-221-527-3
22	106.8	4.5	906	3	US-09-221-527-1
23	106.8	4.5	906	3	US-09-221-236-3
24	106.8	4.5	906	3	US-09-221-416-3
25	106.8	4.5	906	3	US-09-221-245-3
26	106.8	4.5	906	3	US-09-163-115-3
27	106.8	4.5	906	3	US-09-221-528-3

28 106.8 4.5 906 3 US-09-593-553-3 Sequence 3, Appli
29 106.8 4.5 906 3 US-09-221-237-3 Sequence 3, Appli
30 106.8 4.5 906 4 US-09-757-982-3 Sequence 3, Appli
31 106.8 4.5 4137 3 US-09-221-235-1 Sequence 1, Appli
32 106.8 4.5 4137 3 US-09-221-928-1 Sequence 1, Appli
33 106.8 4.5 4137 3 US-09-221-527-1 Sequence 1, Appli
34 106.8 4.5 4137 3 US-09-221-236-1 Sequence 1, Appli
35 106.8 4.5 4137 3 US-09-221-416-1 Sequence 1, Appli
36 106.8 4.5 4137 3 US-09-221-245-1 Sequence 1, Appli
37 106.8 4.5 4137 3 US-09-163-115-1 Sequence 1, Appli
38 106.8 4.5 4137 3 US-09-221-528-1 Sequence 1, Appli
39 106.8 4.5 4137 3 US-09-593-553-1 Sequence 1, Appli
40 106.8 4.5 4137 3 US-09-221-237-1 Sequence 1, Appli
41 106.8 4.5 4137 4 US-09-757-982-1 Sequence 1, Appli
42 100 4.2 1060 4 US-09-579-664B-2 Sequence 2, Appli
43 100 4.2 1060 4 US-10-355-975A-2 Sequence 2, Appli
44 89.6 3.8 7218 1 US-08-232-463-14 Sequence 14, Appli
45 86.2 3.6 4880 3 US-09-031-563-1 Sequence 1, Appli

ALIGNMENTS

RESULT 1

US-09-873-404-1

; Sequence 1, Application US/09873404

; Patent No. 6500656

; GENERAL INFORMATION:

; APPLICANT: WEBSTER, Marion et al

; TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC

; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES

; TITLE OF INVENTION: THEREOF

; FILE REFERENCE: CL001212-CIP

; CURRENT APPLICATION NUMBER: US/09/873,404

; CURRENT FILING DATE: 2001-06-05

; NUMBER OF SEQ ID NOS: 4

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 1

; LENGTH: 2370

; TYPE: DNA

; ORGANISM: Human

; US-09-873-404-1

Query Match 100.0%; Score 2370; DB 4; Length 2370;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2370; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 CATTGGAGACCATGGATAAGTACGATGTGATTAAAGGCCATCGGGCAAGGTGCCTTCGGGA 60
Db 1 CATTGGAGACCATGGATAAGTACGATGTGATTAAAGGCCATCGGGCAAGGTGCCTTCGGGA 60
Qy 61 AAGCATCTTAGCTAAAGGGAAATCAGATAGCAAGCACTGTGTCTATAAAAGAGATCAATT 120
Db 61 AAGCATCTTAGCTAAAGGGAAATCAGATAGCAAGCACTGTGTCTATAAAAGAGATCAATT 120
Qy 121 TTGAAAAGATGCCCATACAGAAAAGAGCTTCAAGAAAAGAGTGAATCTTCTGAAA 180
Db 121 TTGAAAAGATGCCCATACAGAAAAGAGCTTCAAGAAAAGAGTGAATCTTCTGAAA 180
Qy 181 AGATGAACATCCCAACATTTGATGGAGGGATCTCATGAAAAGGATCAATAGACAACGGG 240
Db 181 AGATGAACATCCCAACATTTGATGGAGGGATCTCATGAAAAGGATCAATAGACAACGGG 240
Qy 241 TTATTGTAATGGAAATTTGATGGAGGGATCTCATGAAAAGGATCAATAGACAACGGG 300
Db 241 TTATTGTAATGGAAATTTGATGGAGGGATCTCATGAAAAGGATCAATAGACAACGGG 300
Qy 301 GTGTGTTATTAGTGAAGATCAGATCCTCGGTGGTTGTACAGATTTCTTAGACTAA 360
Db 301 GTGTGTTATTAGTGAAGATCAGATCCTCGGTGGTTGTACAGATTTCTTAGACTAA 360
Qy 361 AACATATTATCATGACAGGAAGATATTACACAGGGACATAAAAGCTCAGAACTTTTCTTA 420

Db 361 AACATATTTCATGACAGGAAGATATTACACAGGACATAAAAGCTCAGAACATTTTCTTA 420
QY 421 GCAAGACGGAATGCTGCAAGAGCTTTGGGTATAGCAAGAGTCTCGAATAATT 480
Db 421 GCAAGACGGAATGCTGCAAGAGCTTTGGGTATAGCAAGAGTCTCGAATAATT 480
QY 481 CCATGGAATCTGCTCGAATCTGTATTTGGAACACCTTACTCTGCTCCCGAGAGATCTGTC 540
Db 481 CCATGGAATCTGCTCGAATCTGTATTTGGAACACCTTACTCTGCTCCCGAGAGATCTGTC 540
QY 541 AGAATAAACCCCTACAAACATAAAGCGGATATTGGTCTCTTGGCTGTGTCTTATATGAGC 600
Db 541 AGAATAAACCCCTACAAACATAAAGCGGATATTGGTCTCTTGGCTGTGTCTTATATGAGC 600
QY 601 TCTGCACACTTAACATCTTTTGGGGTAAACAATTACAGCAGCTGGTCTTGAAGATTT 660
Db 601 TCTGCACACTTAACATCTTTTGGGGTAAACAATTACAGCAGCTGGTCTTGAAGATTT 660
QY 661 GTCAGACACATTTTGGCCCAATATCTCCGGGGTTTCTCGTGAGCTCCATTCCTTGATAT 720
Db 661 GTCAGACACATTTTGGCCCAATATCTCCGGGGTTTCTCGTGAGCTCCATTCCTTGATAT 720
QY 721 CTCAGCTCTTTCAAGTATCTCTCGAGACCGACCATCATAAATTCATTTTGAAGAGC 780
Db 721 CTCAGCTCTTTCAAGTATCTCTCGAGACCGACCATCATAAATTCATTTTGAAGAGC 780
QY 781 CTTTATTAGAAATCTTATTTCCCAATATTGACTCTGAGGTCAATTCAGGAAGATTTCA 840
Db 781 CTTTATTAGAAATCTTATTTCCCAATATTGACTCTGAGGTCAATTCAGGAAGATTTCA 840
QY 841 GTCACATCTTATATGACAGCAGGAGCGCCAGCTTCTCGACATGCTGGGAAGTGGTCC 900
Db 841 GTCACATCTTATATGACAGCAGGAGCGCCAGCTTCTCGACATGCTGGGAAGTGGTCC 900
QY 901 AGAAGTGTAAATCAAAAAGTGAATTCGAGGAAAGTGCCCAACCAAGATCAAGATAT 960
Db 901 AGAAGTGTAAATCAAAAAGTGAATTCGAGGAAAGTGCCCAACCAAGATCAAGATAT 960
QY 961 CTGTGCCAATTTAAAGGAATCTATATTGCAATGAAATGAATGAGAGACCAACAGCTGGAG 1020
Db 961 CTGTGCCAATTTAAAGGAATCTATATTGCAATGAAATGAATGAGAGACCAACAGCTGGAG 1020
QY 1021 CCCAAGAGCCAGATCTATAAAATATAGAAAGACCCAAATTTGCTGTCTGTGGAC 1080
Db 1021 CCCAAGAGCCAGATCTATAAAATATAGAAAGACCCAAATTTGCTGTCTGTGGAC 1080
QY 1081 ATTATGATTATTATGCTCACTTGTATGCTGAGGAGGAGCCCAACCAACCAAGTT 1140
Db 1081 ATTATGATTATTATGCTCACTTGTATGCTGAGGAGGAGCCCAACCAACCAAGTT 1140
QY 1141 ATCAACCTTATTTCTCAAGAAATATCTGGAGTTGAGGATTAACGTTCAAGAAACAGGACATG 1200
Db 1141 ATCAACCTTATTTCTCAAGAAATATCTGGAGTTGAGGATTAACGTTCAAGAAACAGGACATG 1200
QY 1201 GTCCATCCCAAGTCAATGCGCTGCTGAGTACCTTACAGAAATTTGAAGCTCAACAT 1260
Db 1201 GTCCATCCCAAGTCAATGCGCTGCTGAGTACCTTACAGAAATTTGAAGCTCAACAT 1260
QY 1261 ATAAGTTGAAAGTGGAAGCAATTTGGTCTTCTGCTCATCTTCTGCGAGCCCAATTTACA 1320
Db 1261 ATAAGTTGAAAGTGGAAGCAATTTGGTCTTCTGCTCATCTTCTGCGAGCCCAATTTACA 1320
QY 1321 ACCAGACAGAGCTAAGAGTAAATGGAGAGCCTAGATTCCAGGAGCTGCCATTTA 1380
Db 1321 ACCAGACAGAGCTAAGAGTAAATGGAGAGCCTAGATTCCAGGAGCTGCCATTTA 1380
QY 1381 GGAAGACGAAATCAAGGAACAGGAATATTGGAGCAGATTAGAGGAATAACGCCACAGT 1440
Db 1381 GGAAGACGAAATCAAGGAACAGGAATATTGGAGCAGATTAGAGGAATAACGCCACAGT 1440
QY 1441 ACCTCAATGATGAAGAAATTAAGAAAGATTTGGGAGAGAACCCAGAGGACATTTGAAA 1500
Db 1441 ACCTCAATGATGAAGAAATTAAGAAAGATTTGGGAGAGAACCCAGAGGACATTTGAAA 1500

QY 1501 AAGACTTGAAACAAATGAGGCTTCAGAACACAAAGGAAAGTAAAAATCCAGAACAGAAAT 1560
Db 1501 AAGACTTGAAACAAATGAGGCTTCAGAACACAAAGGAAAGTAAAAATCCAGAACAGAAAT 1560
QY 1561 ATAAAGCTTAAGAGGGGTAAATTTGAAATTAATTTAGACAAATGTATTTCTGATGAAA 1620
Db 1561 ATAAAGCTTAAGAGGGGTAAATTTGAAATTAATTTAGACAAATGTATTTCTGATGAAA 1620
QY 1621 ACATCTCTCAAGAGGAAGAGCAATGGAATATACCAAAATGAAACTTTGACCTTTGAGGATG 1680
Db 1621 ACATCTCTCAAGAGGAAGAGCAATGGAATATACCAAAATGAAACTTTGACCTTTGAGGATG 1680
QY 1681 GCATGAAGTTTAAGGAATATGAATGTGTAAGGAGCATGAGATTTATACAGACAAAGCAT 1740
Db 1681 GCATGAAGTTTAAGGAATATGAATGTGTAAGGAGCATGAGATTTATACAGACAAAGCAT 1740
QY 1741 TTGAAAAAATCTTCACTGCCAGAGCAGCATTTACAGAACTGACTTTGGCTCAGTTTCTCT 1800
Db 1741 TTGAAAAAATCTTCACTGCCAGAGCAGCATTTACAGAACTGACTTTGGCTCAGTTTCTCT 1800
QY 1801 TCCTGGAATACTCTCTGCTCATTTCTTCTGAAAAATCTCCATTCAGCAGGATCTTTA 1860
Db 1801 TCCTGGAATACTCTCTGCTCATTTCTTCTGAAAAATCTCCATTCAGCAGGATCTTTA 1860
QY 1861 TTGAGATCTCTCTTGTGCAACGACTCTCTCACTGAGGAGCTGGAGTGAAGGAAATGG 1920
Db 1861 TTGAGATCTCTCTTGTGCAACGACTCTCTCACTGAGGAGCTGGAGTGAAGGAAATGG 1920
QY 1921 AGCTTAGGACATAAACCTTACCACATACATAAAACAAACTTTGGAGAAATCAGGAGAGTAAA 1980
Db 1921 AGCTTAGGACATAAACCTTACCACATACATAAAACAAACTTTGGAGAAATCAGGAGAGTAAA 1980
QY 1981 GCCAAGAGGAGAGACAGAGTCAATGGGAGGACACAGAAATTTGGCAGCATCAACTGGGAAGA 2040
Db 1981 GCCAAGAGGAGAGACAGAGTCAATGGGAGGACACAGAAATTTGGCAGCATCAACTGGGAAGA 2040
QY 2041 GAAAGCCAGATGAGGGTTTCCAGCGAGACTGTAGTTGCTGTGGGAAAACAGGAGGAGT 2100
Db 2041 GAAAGCCAGATGAGGGTTTCCAGCGAGACTGTAGTTGCTGTGGGAAAACAGGAGGAGT 2100
QY 2101 GGGATGGAGGAGCGCTCAGACTCTGTGCAAGATGATGGCAGTGGCCGACATCACTCCA 2160
Db 2101 GGGATGGAGGAGCGCTCAGACTCTGTGCAAGATGATGGCAGTGGCCGACATCACTCCA 2160
QY 2161 CCGTCCCAAGCGGCTGACAGTGTGCTGCTTGTAGTGTGCTGAGTGTGCTGAGGAGGAG 2220
Db 2161 CCGTCCCAAGCGGCTGACAGTGTGCTGCTTGTAGTGTGCTGAGTGTGCTGAGGAGGAG 2220
QY 2221 CCAAGGACCCGTACAGCCAGTGTCTCATCTGTGATGTGATGTCTTCTTCTACCTATACAC 2280
Db 2221 CCAAGGACCCGTACAGCCAGTGTCTCATCTGTGATGTGATGTCTTCTTCTACCTATACAC 2280
QY 2281 CCTATAGATCTGTATCAGACACTTTCAATATGCTTTTGTATATCTCAGAAAAAAA 2340
Db 2281 CCTATAGATCTGTATCAGACACTTTCAATATGCTTTTGTATATCTCAGAAAAAAA 2340
QY 2341 AAAAAAACACTGTCTCATGCCGTACGAGCG 2370
Db 2341 AAAAAAACACTGTCTCATGCCGTACGAGCG 2370

RESULT 2

US-10-243-735-1
; Sequence 1, Application US/10243735
; Patent No. 6706510
; GENERAL INFORMATION:
; APPLICANT: WEBSTER, Marion et al
; TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC
; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES
; FILE REFERENCE: CL001212DIV
; CURRENT APPLICATION NUMBER: US/10/243,735

; CURRENT FILING DATE: 2002-09-16									
; NUMBER OF SEQ ID NOS: 4									
; SOFTWARE: FastSeq for Windows Version 4.0									
; SEQ ID NO 1									
; LENGTH: 2370									
; TYPE: DNA									
; ORGANISM: Human									
US-10-243-735-1									
Query Match									
Best Local Similarity 100.0%; Score 2370; DB 4; Length 2370;									
Matches 2370; Conservative 0; Mismatches 0; Indels 0; Gaps 0;									
Qy	1	CATTGGAGACCATGGTAAAGTACGATGCTGATTAAGCCATCGGCAAGTGCCTTCGGGA	60						
Db	1	CATTGGAGACCATGGTAAAGTACGATGCTGATTAAGCCATCGGCAAGTGCCTTCGGGA	60						
Qy	61	AAGCATCTAGTCTAAAGGGAATCAGATAGCAAGCACTGTCTATAAAGAGATCAATT	120						
Db	61	AAGCATCTAGTCTAAAGGGAATCAGATAGCAAGCACTGTCTATAAAGAGATCAATT	120						
Qy	121	TTGAAAGATGCCATACAAAGAAAGAGCTTCAAAGAAAGATGATTTCTTGGAAA	180						
Db	121	TTGAAAGATGCCATACAAAGAAAGAGCTTCAAAGAAAGATGATTTCTTCTGGAAA	180						
Qy	181	AGATGAACATCCCAACTGTAGCTTCTTCAATTCTTCAAGAGATGCGAGCTGT	240						
Db	181	AGATGAACATCCCAACTGTAGCTTCTTCAATTCTTCAAGAGATGCGAGCTGT	240						
Qy	241	TTATTGTAATGGAATTTGTATGAGGGGATCTCATGAAAGGATCAATAGACAACGG	300						
Db	241	TTATTGTAATGGAATTTGTATGAGGGGATCTCATGAAAGGATCAATAGACAACGG	300						
Qy	301	GTGTGTTATTAGTGAAGATCAGATCCCTCGTGTGTTGTACAGATTTCTTAGACTAA	360						
Db	301	GTGTGTTATTAGTGAAGATCAGATCCCTCGTGTGTTGTACAGATTTCTTAGACTAA	360						
Qy	361	AAACATATTCATGACAGGAAGATATTACACAGGACATAAAGCTCAGAACATTTTCTTA	420						
Db	361	AAACATATTCATGACAGGAAGATATTACACAGGACATAAAGCTCAGAACATTTTCTTA	420						
Qy	421	GCAAGAACGGATGTGGCAAGCTTGGGACTTTGGTATAGCAAGATCTCTGAATTAAT	480						
Db	421	GCAAGAACGGATGTGGCAAGCTTGGGACTTTGGTATAGCAAGATCTCTGAATTAAT	480						
Qy	481	CCATGGAATCTGCTGAACTTGTATTGGAACACCTTACTACCTGCCCCAGAGATCTGTC	540						
Db	481	CCATGGAATCTGCTGAACTTGTATTGGAACACCTTACTACCTGCCCCAGAGATCTGTC	540						
Qy	541	AGATTAACCCCTACAACTAAACCGATATTGCTGCTTGGCTGTGCTTATATGAGC	600						
Db	541	AGATTAACCCCTACAACTAAACCGATATTGCTGCTTGGCTGTGCTTATATGAGC	600						
Qy	601	TCTGCACTTAAACATCTTTTGAGGGTAAACCTTACAGCAGCTGCTTCTGAGATTT	660						
Db	601	TCTGCACTTAAACATCTTTTGAGGGTAAACCTTACAGCAGCTGCTTCTGAGATTT	660						
Qy	661	GTCAAGCACATTTTGCCCCAATATCTCCGGGGTTTTCTCGTGAGCTCCATTCCTTGAT	720						
Db	661	GTCAAGCACATTTTGCCCCAATATCTCCGGGGTTTTCTCGTGAGCTCCATTCCTTGAT	720						
Qy	721	CTCAGCTCTTTCAAGTATCTCTCGAGCCGACCATCAATAATTCATTTGAAAGGC	780						
Db	721	CTCAGCTCTTTCAAGTATCTCTCGAGCCGACCATCAATAATTCATTTGAAAGGC	780						
Qy	781	CTTTTGTAGAGATCTTATTCCCAATATTGACTCTTGAGGTCAATCAGGAAGAAATCA	840						
Db	781	CTTTTGTAGAGATCTTATTCCCAATATTGACTCTTGAGGTCAATCAGGAAGAAATCA	840						
Qy	841	GTCAATGCTTTATGACAGAGAGAGCGCCAGCTTCTGCAATGCTGGGAAGGTGCTCC	900						
Db	841	GTCAATGCTTTATGACAGAGAGAGCGCCAGCTTCTGCAATGCTGGGAAGGTGCTCC	900						

Qy	901	AGAAAGTGTAAAAATACAAAAAGTGAATTCGGGGAAGTCCCCACCAAGATCAAGGATAT	960						
Db	901	AGAAAGTGTAAAAATACAAAAAGTGAATTCGGGGAAGTCCCCACCAAGATCAAGGATAT	960						
Qy	961	CTGTGCCAATTTAAAGGAATGCTATATTGTCATAGAAATGAATGGAGACCAAGCTGGAG	1020						
Db	961	CTGTGCCAATTTAAAGGAATGCTATATTGTCATAGAAATGAATGGAGACCAAGCTGGAG	1020						
Qy	1021	CCAGAAAGCCAGATCTATAAAAAATGATAGAAAGACCAAAAAATGCTGTCTGGAC	1080						
Db	1021	CCAGAAAGCCAGATCTATAAAAAATGATAGAAAGACCAAAAAATGCTGTCTGGAC	1080						
Qy	1081	ATTATGATTATTATTATGCTCAACTTGTATGCTGAGGAGGAGCCCAAAACCAAGTT	1140						
Db	1081	ATTATGATTATTATTATGCTCAACTTGTATGCTGAGGAGGAGCCCAAAACCAAGTT	1140						
Qy	1141	ATCACCCCTATTCTCAAGAAAAATCTGGAGTTTACGGTTCAGGAAACGAGGCATG	1200						
Db	1141	ATCACCCCTATTCTCAAGAAAAATCTGGAGTTTACGGTTCAGGAAACGAGGCATG	1200						
Qy	1201	GTCCATCCCAAGTCANTGCCCTGCTGAGTACCTTCAGAGAAAAATTTGAGCTCAACAT	1260						
Db	1201	GTCCATCCCAAGTCANTGCCCTGCTGAGTACCTTCAGAGAAAAATTTGAGCTCAACAT	1260						
Qy	1261	ATAAGTTGAAAGTGGAGAACAAATTTGGTCTTCTGTCATCTTCTGCGGACCAAAATACA	1320						
Db	1261	ATAAGTTGAAAGTGGAGAACAAATTTGGTCTTCTGTCATCTTCTGCGGACCAAAATACA	1320						
Qy	1321	ACGAGACAAGAGCTTAAGAAATTAAGGAGAGAGCTAGATTCCTAGGAGCTGCCATTTA	1380						
Db	1321	ACGAGACAAGAGCTTAAGAAATTAAGGAGAGAGCTAGATTCCTAGGAGCTGCCATTTA	1380						
Qy	1381	GGAAAAACGAAATCAAGGAACAGGAATATTGGAGAGAGTATAGAGGAAATACGCCAACAGT	1440						
Db	1381	GGAAAAACGAAATCAAGGAACAGGAATATTGGAGAGAGTATAGAGGAAATACGCCAACAGT	1440						
Qy	1441	ACCTCAATGATCAAAAGAAAAATTTAGAAAGAAAGTGGGAGAGAACACAGAGCAATTGAAA	1500						
Db	1441	ACCTCAATGATCAAAAGAAAAATTTAGAAAGAAAGTGGGAGAGAACACAGAGCAATTGAAA	1500						
Qy	1501	AGAGCTTGAACAAATAGAGCTTCAGAACACAAAGGAAAGTAAATTCAGAGACAGAAAT	1560						
Db	1501	AGAGCTTGAACAAATAGAGCTTCAGAACACAAAGGAAAGTAAATTCAGAGACAGAAAT	1560						
Qy	1561	ATAAAGCTTAAGAGGGGTAAAAATTTGAAATTAATTTAGACAAATGTATTCTTGATGAAA	1620						
Db	1561	ATAAAGCTTAAGAGGGGTAAAAATTTGAAATTAATTTAGACAAATGTATTCTTGATGAAA	1620						
Qy	1621	ACATCTCCAAAGAGAGAGGCAATGGATATACAAATGAAACTTTGACCTTTGAGGATG	1680						
Db	1621	ACATCTCCAAAGAGAGAGGCAATGGATATACAAATGAAACTTTGACCTTTGAGGATG	1680						
Qy	1681	GCATGAAGTTTAAAGGAATATGAAATGTAAGGAGCATGGAGATTATACAGACAAAGCAT	1740						
Db	1681	GCATGAAGTTTAAAGGAATATGAAATGTAAGGAGCATGGAGATTATACAGACAAAGCAT	1740						
Qy	1741	TTGAAAACTTCACTGCCAGAGAGCAATTTACAGAACTGACTTGCTTGCTCTCTCT	1800						
Db	1741	TTGAAAACTTCACTGCCAGAGAGCAATTTACAGAACTGACTTGCTTGCTCTCTCTCT	1800						
Qy	1801	TCCTGGAACTACTCTCTGCCCTCATTTCTCTTGGAAAAATCTCCATTCAGCAGGCATCTTA	1860						
Db	1801	TCCTGGAACTACTCTCTGCCCTCATTTCTCTTGGAAAAATCTCCATTCAGCAGGCATCTTA	1860						
Qy	1861	TTGAGGATCTCTTTGTGCCAACGACTGCTCACTGAAGGAGCTGGAGTGAAGAAATGG	1920						
Db	1861	TTGAGGATCTCTTTGTGCCAACGACTGCTCACTGAAGGAGCTGGAGTGAAGAAATGG	1920						
Qy	1921	AGCTTAGGACATTAACCCCTACCTACATATAAACAATTTTGGAAATCAGGAGAGTAAA	1980						
Db	1921	AGCTTAGGACATTAACCCCTACCTACATATAAACAATTTTGGAAATCAGGAGAGTAAA	1980						
Qy	1981	GCCAAAGGAGGAGAGACAGGTCATGGGAGGACAGGAATTTGGCAGCATCAACTGGGAAGA	2040						

Db 1981 GCCAAAGGAGGAGAGACAGCTCATGGGAGGACACAGGAATGGCAGCATCAACTGGAAGA 2040
QY 2041 GAAAGGCCAGATGAGGGTTTTCCACGACAGACTGATGTTGCTGTGGGAAACAGAGGCGAGT 2100
Db 2041 GAAAGGCCAGATGAGGGTTTTCCACGACAGACTGATGTTGCTGTGGGAAACAGAGGCGAGT 2100
QY 2101 GGGATGAGGAGGCCCTCAGACTCTGCTGCAGATGATGGCGAGTGGCGGACATCACCTCCA 2160
Db 2101 GGGATGAGGAGGCCCTCAGACTCTGCTGCAGATGATGGCGAGTGGCGGACATCACCTCCA 2160
QY 2161 CCTGCCCCACGGGCGCTGACAGTGAAGTCTGTGCTTACTGTGTCAGTGTGTCAGGAAGGGAAGA 2220
Db 2161 CCTGCCCCACGGGCGCTGACAGTGAAGTCTGTGCTTACTGTGTCAGTGTGTCAGGAAGGGAAGA 2220
QY 2221 CCAAGGACCGGTACAGCCCGAGTCTCATCTGATGATAGTCTACTTCTCAGTATACAC 2280
Db 2221 CCAAGGACCGGTACAGCCCGAGTCTCATCTGATGATAGTCTACTTCTCAGTATACAC 2280
QY 2281 CCTATAGATCTGTATCAGACACTTTCAAAATATGTTTGTATATCTCAAGAAAAAAA 2340
Db 2281 CCTATAGATCTGTATCAGACACTTTCAAAATATGTTTGTATATCTCAAGAAAAAAA 2340
QY 2341 AAAAAAACACTGTCATGCCGTTACGAGCG 2370
Db 2341 AAAAAAACACTGTCATGCCGTTACGAGCG 2370

RESULT 3

US-09-620-312D-246

; Sequence 246, Application US/09620312D

; Patent No. 6569662

; GENERAL INFORMATION:

; APPLICANT: Tang, Y. Tom

; APPLICANT: Liu, Chenghua

; APPLICANT: Asundi, Vinod

; APPLICANT: Zhang, Jie

; APPLICANT: Ren, Feiyan

; APPLICANT: Chen, Rui-hong

; APPLICANT: Zhao, Qing A.

; APPLICANT: Wehrman, Tom

; APPLICANT: Xue, Aidong J.

; APPLICANT: Yang, Yonghong

; APPLICANT: Wang, Jian-Rui

; APPLICANT: Zhou, Ping

; APPLICANT: Ma, Yunqing

; APPLICANT: Wang, Dunrui

; APPLICANT: Wang, Zhiwei

; APPLICANT: John Tillinghast

; APPLICANT: Drmanac, Radoje T.

; TITLE OF INVENTION: No. 6569662el Nucleic Acids and

; TITLE OF INVENTION: Polypeptides

; FILE REFERENCE: 784C1P2B

; CURRENT APPLICATION NUMBER: US/09/620,312D

; CURRENT FILING DATE: 2000-07-19

; PRIOR APPLICATION NUMBER: 09/552,317

; PRIOR FILING DATE: 2000-04-25

; PRIOR APPLICATION NUMBER: 09/488,725

; PRIOR FILING DATE: 2000-01-21

; NUMBER OF SEQ ID NOS: 1105

; SOFTWARE: Pt_FL_Genes Version 1.0

; SEQ ID NO 246

; LENGTH: 5448

; TYPE: DNA

; ORGANISM: Homo sapiens

; FEATURE:

; NAME/KEY: CDS

; LOCATION: (534)..(4178)

; US-09-620-312D-246

Query Match

Best Local Similarity 15.6%; Score 369.4; DB 4; Length 5448;

Matches 551; Conservativity 0; Mismatches 281; Indels 3; Gaps 1;

QY 9 ACCATGGATAAGTACGATGTGATTAAAGCCATCGGCAAGGTGCTTCGGGAAAGCATAC 68
Db 531 ATCATGGAGAGTATGTTAGACTACAGAGATTCGAGAGGTTTCATTTGGAAAGCCATT 590
QY 69 TTAGCTAAAGGGAATCAGATAGCAAGCACGTGTCTATAAAAGAGATCAATTTTGAAGAAG 128
Db 591 CTTGTAAATCTACAGAGATGGCAGACAGTATGTTATCAAGGAAATTAACATCTCAAGA 650
QY 129 ATGCCCATACAGAAAGAGAGCTTCAAGAAAGAGTGTCTTCTGGAAGATGAAA 188
Db 651 ATGTCCAGTAAGAAAGAGAAAGTCAAGGAGAGAGTTCGAGTATTGGCAACATGAAG 710
QY 189 CATCCCAACATGTAGCCCTTCTCAATTTCAATTTCAAGAGAAATGCGAGGCTGTTTATGTA 248
Db 711 CATCAAAATATGTCAGATATAGAGAAATCATTTGAAGAAATGCTCTCTACATAGTA 770
QY 249 ATGGAATATTGTATGGAGGAGTCTCATGAAAGGATCAATAGACAACGGGTGTCGTA 308
Db 771 ATGGAATATTGTATGGAGGAGTCTCTGTTAAAGCAATAAATGCTCAGAAAGGCTTTG 830
QY 309 TTTAGTGAAGATCAGATCCTCGGTTGGTTTGTACAGATTTCTCTAGGACTAAACATATT 368
Db 831 TTTCAAGAGGATCAGATTTTGGACTGGTTTGTACAGATATGTTGGCCCTGAAACATGTA 890
QY 369 CATGACGAGAGATATTACACAGGAGACATAAAAGCTCAGAACATTTTCTTAGCAAGAAC 428
Db 891 CATGATAGAAAATTTCTTCATCGAGACATTAATCTCAGAACATATTTTAACTAAAGAT 950
QY 429 GGAATGTGGCAAGAGCTTTGGGAGCTTTGGTATAGCAAGAGTCTGTAATATTCATCGAA 488
Db 951 GGAACAGT---ACAACTTGGAGATTTTGGATTCCTAGAGTCTTAATAGTACTGTAGAG 1007
QY 489 CTTGCTCGAATTTGTATTGGAACACCTTACTACTCTGCTGCCAGAGATCTGTCAAGATAAA 548
Db 1008 CTGGCTCGAATCTGCATAGAGGAGCCCATACTACTTGTCACTGAAATCTGTGAAAAAATA 1057
QY 549 CCCTACACATTAACCGGATATTTGGTCTCTTGGCTGTCTTATATAGCTCTGCACA 608
Db 1068 CTTTACAATAATAAAGTGACATTTGGGCTCTGGGCTGTGCTTTTATGAGCTGTGACA 1127
QY 609 CTTAAACATCTTTTGGGGTAAACACTTACAGCAGCTGTTCTGAAGATTTGTCAAGCA 668
Db 1128 CTTAAACATGCTTTTGAAGCTGGCAGTATGAAAAAAGCTGCTGAGAGATAATATCTGGA 1187
QY 669 CATTTTGGCCCAATATCTCCGGGTTTTTCTCGTGAAGCTCCATCTCTGATATCTCAGCTC 728
Db 1188 TCTTTTCCACCTGTGCTTTGCTATTTCCCTATGATCTCCGAGTTTGGTGTCTCAGTTA 1247
QY 729 TTTCAAGTATCTCTCGAGACCGGACCATCCATAAATTTCCATTTTGAAGGCGCTTTTATA 788
Db 1248 TTTAAAGAAATCTAGGATAGACCATCAGTCACTCCATATTTGGAGAAAGTTTATA 1307
QY 789 GAGAACTTATTTCCAAATATTTCAGTCTCCGAGTCAATTCAGGAAGATTCAGTCTC 843
Db 1308 GCCAAACGATTGAAAGTTTCTCTCTCAGCTTATTCAGAAAGATTTTGTCTC 1362

RESULT 4

US-09-620-312D-245

; Sequence 245, Application US/09620312D

; Patent No. 6569662

; GENERAL INFORMATION:

; APPLICANT: Tang, Y. Tom

; APPLICANT: Liu, Chenghua

; APPLICANT: Asundi, Vinod

; APPLICANT: Zhang, Jie

; APPLICANT: Ren, Feiyan

; APPLICANT: Chen, Rui-hong

; APPLICANT: Zhao, Qing A.

; APPLICANT: Wehrman, Tom

; APPLICANT: Xue, Aidong J.

; APPLICANT: Yang, Yonghong

```
; APPLICANT: Wang, Jian-Rui
; APPLICANT: Zhou, Ping
; APPLICANT: Ma, Yungling
; APPLICANT: Wang, Dunrui
; APPLICANT: Wang, Zhiwei
; APPLICANT: John Tillinghast
; APPLICANT: Dmanac, Radoje T.
; TITLE OF INVENTION: No. 6569662el Nucleic Acids and
; FILE REFERENCE: Polypeptides
; CURRENT APPLICATION NUMBER: US/09/620,312D
; PRIOR FILING DATE: 2000-07-19
; PRIOR APPLICATION NUMBER: 09/552,317
; PRIOR FILING DATE: 2000-04-25
; PRIOR APPLICATION NUMBER: 09/488,725
; PRIOR FILING DATE: 2000-01-21
; NUMBER OF SEQ ID NOS: 1105
; SOFTWARE: pt_PL_genes Version 1.0
; SEQ ID NO 245
; LENGTH: 5532
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (534)..(4262)
; US-09-620-312D-245

Query Match 15.6%; Score 369.4; DB 4; Length 5532;
Best Local Similarity 66.0%; Pred. No. 1.5e-97;
Matches 551; Conservative 0; Mismatches 281; Indels 3; Gaps 1;

QY 9 ACCATGGTAAGTACATGCTGATTAGGCCATCGGCCAGGTGCTTCGGGAAGCATAC 68
DB 531 ATCATGGGAAGTATGTTAGACTACAGAAGATTGGAGAAGTTCATTGGAAAGCCATT 590
QY 69 TTAGCTAAAGGGAAATCAGATAGCAAGCACTGTGTCTATAAAAGAGATCAATTTTGAAG 128
DB 591 CTGTGTAATCTACAGAAGATGCGACAGATATGTTATCAAGAAATTAACATCTCAGA 650
QY 129 ATGCCCATACAGAAAAGAGCTTCAAGAAAGAGTGAATCTCTGGAAAAGATGAAA 188
DB 651 ATGTCCAGTAAAGAAAGAGAAGAAATCAAGAGAGAGTTCAGTATATGGCAACATGAAG 710
QY 189 CATCCCAACATTGTAGCTTCTTCAATTCATTTCAAGAAAGTGGAGGCTGTTTATTGTA 248
DB 711 CATCCAAATATGTCAGATATAGAGATCATTTTGAAGAAATGGCTCTCTACATAGTA 770
QY 249 ATGGAATATTTGATCGAGGGGATCTCATGAAAGAGATCAATAGACAACGGGCTGTGTA 308
DB 771 ATGGATTACTGTGAGGGAGGGATCTGTTTAAAGCGAATAAATGCTCAGAAAAGCGTTTG 830
QY 309 TTTAGTGAAGATCAGATCTCGTGGTTTGTACAGATTTCTTAGGACTAAACATATT 368
DB 831 TTTCAAGAGATCAGATTTGGACTGGTTTGTACAGATATGTTGGCCCTGAAACATGTA 890
QY 369 CATGACAGGAAGATATTACACAGGGACATATAAGCTCAGAACATTTTCTTAGCAAGAAC 428
DB 891 CATGATAGAAAATTTCTTCATCGAGACATTAATCTCAGACATATTTTAACTAAAGAT 950
QY 429 GGAATGGTGGAAAGCTTGGGACTTTGGTATAGCAGAGTCTCGAATATTCATTCGAA 488
DB 951 GGAACAGT---ACAACCTCGGATTTTGGAAATTTGCTAGAGTTCTTAATAGTACTGTAG 1007
QY 489 CTGTGCTCGAACTGTATTGGAACACCTTACTACCTGCTCCAGAGATCTGTGAGATAAA 548
DB 1008 CTGGCTCGAACTGTGATAGGACCCCATACTACTGTACCTGGAATCTGTGAAACAAA 1067
QY 549 CCTTACAAACATAAAACGGATATTTGGTCTCTTGGCTGTGTCTTATATGAGCTCTGACA 608
DB 1068 CCTTACAAATAAATAAGTGACATTTGGGCTCTGGGCTGTGCTCTTATGAGCTGTGTACA 1127
QY 609 CTTAAACATCCTTTTCAGGGTAAACATTTACAGCAGCTGTTCTGAGAGTTTGTCAAGCA 668
```

```
DB 1128 CTTAAACATGCTTTTGAAGCTGGCAGTATGAAAAAACCCTGCTACTGAAGATAATATCTGGA 1187
QY 669 CATTTTGCCTCAATATCTCGGGGTTTCTCGTGAGCTCCATTCCTTGATATCTCAGCTC 728
DB 1188 TCTTTTCCACTGTGTCTTTGCAATATCTCTATGATCTCCGAGTTTGGTGTCTCAGTTA 1247
QY 729 TTTCAAGTATCTCTCGAGACCGGACCATCCATATAATTTCCATTTTGAAGAGCCCTTTTA 788
DB 1248 TTTAAAGAAATCTCAGGATAGACCATCAGTCAATCCATATTTGGAGAAAGGTTTATA 1307
QY 789 GAGAACTTTATTTCCCAATATTTGACCTCCGAGGTCTATTCAGGAAGAAATTCAGTC 843
DB 1308 GCCAAACGCATTGAAAGTTTCTCTCTCTCAGCTTATTCGAGAAGAAATTTTGTG 1362

RESULT 5
US-09-166-350-29
; Sequence 29, Application US/09166350A
; Patent No. 6440863
; GENERAL INFORMATION:
; APPLICANT: Scanlan, Matthew
; APPLICANT: Chen, Yao
; APPLICANT: Stockert, Elisabeth
; APPLICANT: Old, Lloyd
; APPLICANT: Jager, Elke
; APPLICANT: Knuth, Alex
; TITLE OF INVENTION: Renal Cancer Associated Antigens and
; FILE REFERENCE: L0461/7051
; CURRENT APPLICATION NUMBER: US/09/166,350A
; CURRENT FILING DATE: 1998-10-05
; EARLIER APPLICATION NUMBER: US 09/166,350
; EARLIER FILING DATE: 1998-10-05
; NUMBER OF SEQ ID NOS: 35
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 29
; LENGTH: 4263
; TYPE: DNA
; ORGANISM: Homo sapiens
; US-09-166-350-29

Query Match 15.0%; Score 355.8; DB 4; Length 4263;
Best Local Similarity 64.8%; Pred. No. 1.3e-93;
Matches 544; Conservative 0; Mismatches 292; Indels 3; Gaps 1;

QY 5 GGAGACCATGGATGAAGTACGATGTGATTAGGCCATCGGCAAGGTCCCTTCGGGAAGC 64
DB 569 GGGAAACCATGGAGAAGTATGTGAGACTGCAGAAAGTTGGAAGAGTTCAATTTGGAAAAGC 628
QY 65 ATACTTAGCTAAAGGGAAATCAGATAGCAAGCACTGTGTCTATAAAAGAGATCAATTTTGA 124
DB 629 TGTTCCTTGTAAATCGACAGAGGATGCGACATATTATGTCATCAAGGAATTAACATCTC 688
QY 125 AAAGATGCCCATACAGAAAAGAGCTTCAAGAAAGAGTGAATTTCTCTGGAAGAAAT 184
DB 689 AAGAATCTCTGATTAAGAAAGGCAAGAAATCAAGAGAGAGTTCCTGTATTGGCAACAT 748
QY 185 GAAACATCCCAACATTTGAGCTTCTTCAATTTCAAGAGAGTGGAGGCTGTTTAT 244
DB 749 GAACATCCCAATATTTGTCATATAAAGAAATCATTGGAAGAAATGCTCTCTCTACAT 808
QY 245 TGTAAATGGAATTTTGTGATGGAGGGATCTCATGAAAGAGATCAATAGACAACGGGCTGT 304
DB 809 AGTAATGGATTTCTGTCAAGAGGCTGATTTGTTTAAACGAATAAATGCTCAGAAAGCGC 868
QY 305 GTTATTTAGTGAAGATCAGATCTCGGTTGGTTTGTACAGATTTCTTAGACTTAAACA 364
DB 869 TCTGTTTCAAGAGACAGATTTTGGACTGGTTTGTGCAATATGTTTGGCTCTGGAAGCA 928
QY 365 TATTTCATGACAGGAAGATATTACACAGGGACATATAAGAGCTCAGAAACATTTTCTTAGCAA 424
DB 929 TGTACATGATAGAAAATTTCTTCAACCGAGACATAAAGTCCACAGAACATATTTTCTAACCA 988
```

Qy	425	GAACGGAAATG	TGCGCAAGCTT	GGGACATTT	GGGTATAGCAAGATCCT	CGAATAATTCAT	484
Db	989	AGATGGGACAGTG	--	-CAGCTTGAGAGATTTT	GGAAATTCGTCGAGTCTCTT	TATAGTACTGT	1045
Qy	485	GGAACTTGC	TGCAACTGTATT	GGAAACACCTTACT	AGCTGCCAGAGATCTCT	CAGAA	544
Db	1046	AGAGCTGGCT	CGAACTTGCA	TAGGCATCCAT	TACTTGTGCACCTGAAATCT	GTGAAA	1105
Qy	545	TAAACCC	TCAACCAATAAAAA	CGGATATTT	TGGTCTCTTGGCTGTCTCTT	TATATGAGCTCTG	604
Db	1106	CAAGCC	TTATPACCAATAA	AGTGACATTTGG	GCCTTGGCGTGTCTCTT	TATAGATTTGTG	1165
Qy	605	CACACTTAA	ACATCCTTTT	TGAGGTACAACATTA	CAGAGCTGTTCTG	GAAGATTTGTCA	664
Db	1166	TACACTTAA	ACATGCATTTGA	AGCTGGAACATGAAAA	CCCTGTA	CTGAAGATAATCTC	1225
Qy	665	AGCACATTTT	CCCCAATATCT	CGGGGTTTTCT	CGTAGCTCCATTTCC	CTTGATATCTCA	724
Db	1226	CGGATCCTT	CTCC	TCCAGTGCTCCACATTA	CTCTCTATGATCTCCG	AGCTTGTGTCTCA	1285
Qy	725	GCTCTTT	CAAGTATCTCCT	CGAGACCGACCATCC	ATAATTTCCATTTT	TGAAAAGCCCTT	784
Db	1286	GTATTTTAA	AGAAATCCTAGG	ATAGACCATCAGTCA	ATCCATATTTGG	GAAGGTTT	1345
Qy	785	TTTAGAAATCT	TTATCCCAAATTT	TGACTCTGAGGT	CATT	CAGGAAGATTTCA	843
Db	1346	TATAGCTTAA	ACGAATCGAAAGTTCT	CTCCCTCAGCTTAT	TTCGAA	GAATTTGTG	1404

RESULT 6

```

US-09-173-581-15
; Sequence 15, Application US/09173581A
; Patent No. 6013455
; GENERAL INFORMATION:
; APPLICANT: Bandman, Olga
; APPLICANT: Tang, Y. Tom
; APPLICANT: Hillman, Jennifer L.
; APPLICANT: Yue, Henry
; APPLICANT: Guegler, Karl J.
; APPLICANT: Corley, Neil C.
; APPLICANT: Gorgone, Gina
; APPLICANT: Azimzai, Yalda
; APPLICANT: Lu, Aina
; TITLE OF INVENTION: Protein Kinase Homologs
; FILE REFERENCE: PF-0614 US
; CURRENT APPLICATION NUMBER: US/09/173,581A
; CURRENT FILING DATE: 1998-10-15
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: PERL Program
; SEQ ID NO 15
; LENGTH: 1846
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE: -
; OTHER INFORMATION: 1567782
US-09-173-581-15

```

RESULT 7

```

US-09-420-915-15
; Sequence 15, Application US/09420915
; Patent No. 626497
; GENERAL INFORMATION:
; APPLICANT: Bandman, Olga
; APPLICANT: Tang, Y. Tom
; APPLICANT: Hillman, Jennifer L.
; APPLICANT: Yue, Henry
; APPLICANT: Guegler, Karl J.
; APPLICANT: Corley, Neil C.
; APPLICANT: Gorgone, Gina
; APPLICANT: Azimzai, Yalda
; APPLICANT: Lu, Aina
; TITLE OF INVENTION: Protein Kinase Homologs
; FILE REFERENCE: PF-0614 US
; CURRENT APPLICATION NUMBER: US/09/420,915
; CURRENT FILING DATE: 1999-10-20
; EARLIER APPLICATION NUMBER: US 09/173,581
; EARLIER FILING DATE: 1998-10-15
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: PERL Program
; SEQ ID NO 15
; LENGTH: 1846
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE: -
; OTHER INFORMATION: 1567782
US-09-420-915-15

```

504	Qy	ATTGGAACACCTTACTACTGCTGCCAGAGATCTGTGAGATTAACCCCTCAACAAATAAA	553
547	Db	ATAGGAGCCCCATCTACTCTTGTCACCTGAATCTGTGAAACCAACCTTTACAATAATAAA	606
564	Qy	ACGGATATTGGTCTCTCTGGGTGTGCTTTATATAGAGCTCTGCACATTAACACATCCTTTT	623
607	Db	AGTGACATTTGGGCTCTGGGTGTGTCCTTTATCAGCTGTGTACACATTAACATGCTTTT	666
624	Qy	GAGGGTAAACCTTACACGACCTGGTCTCGAAGATTGTCAAGCACATTTTGGCCCCAATA	683
667	Db	GAACTGGCAGTATGAATAAACCCTGACTGAAGATAAATATCTGGATCTTTTCCACCTGIG	726
684	Qy	TCTCGGGGTTTTCTCTGTGAGCTCCATTCCTTGATATCTCAGCTCTTTCAAGTATCTCCT	743
727	Db	TCTTTGCATTATTCTCTATGATCTCGCAGTTTGGTGTCTCAGTTATTATAAAGAAATCCT	786
744	Qy	CGAGACCCACCATCATTAATTCATTTGAAAGGCCCTTTTAGAGATCTTATTCC	803
787	Db	AGGGATAGACCATCAGTCACTTCATATTGGAAGAAGTTTTATGCCAACGCATTGAA	846
804	Qy	AAATATTGACCTCTGAGGTCATTTCAGGAAGAAATTCAGTC	843
847	Db	AAAGTTCTCTCTCTCCTCAGCTTATTGCAAGAAATTTTGTG	886

Db 547 ATAGGACCCCATACTACTTGTACCTGAAATCTGTGAAACAAACCTTACATAATAA 606
Qy 564 ACGGATATTTGGTCTCTTGGCTGTGCTTATATGAGCTCTGCACATTAACATCTCTTT 623
Db 607 AGTGACATTTGGGCTCTGGGCTGTCTCTTATGAGCTGTGTACACTTAAACATGCTTT 666
Qy 624 GAGGTAACAACTTACAGCAGCTGGTTCTCAAGATTGTCAAGCACAATTTGCCCCAATA 683
Db 667 GAAGCTGGCAGTATGAACAACTGGTACTGAAGATAATATCTGATCTTTCCACCTGTG 726
Qy 684 TCTCCGGGTTTCTCTGAGCTCCATTCCTTGTATATCTCAGCTCTTTCAAGTATCTCT 743
Db 727 TCTTTGCAATATCTTATGATCTCCGAGTTTGGTGTCTCAGTTATTTAAAGAAATCCT 786
Qy 744 CGAGACCGACCAATCAATTAATTCATTTGAAAGGCCCTTTTAGAGATCTTATCC 803
Db 787 AGGATAGACCATCAGTCAACTCCATATTCGAGAAAGGTTTATAGCCAAACGATTGAA 846
Qy 804 AATATTGACTCTCTGAGGTCATTTCAGGAAGAATTCAGTC 843
Db 847 AAGTTCTCTCTCTCAGCTATTTCGAGAAGAATTTTGIC 886

RESULT 8
US-09-873-404-3
; Sequence 3, Application US/09873404
; Patent No. 650656
; GENERAL INFORMATION:
; APPLICANT: WEBSTER, Marion et al
; TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC
; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES
; FILE REFERENCE: CL001212-CIP
; CURRENT APPLICATION NUMBER: US/09/873,404
; CURRENT FILING DATE: 2001-06-05
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 63588
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(63588)
; OTHER INFORMATION: n = A,T,C or G
US-09-873-404-3

Query Match 8.7%; Score 205.2; DB 4; Length 63588;
Best Local Similarity 96.3%; Pred. No. 1.2e-48;
Matches 210; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

Qy 617 TCCTTTTGAGGGTAACTTACAGCAGCTGGTTCTGAAAGATTGTCAAGCACAATTTTGC 676
Db 27430 TCAGTTTGAGGGTAACTTACAGCAGCTGGTTCTGAAAGATTGTCAAGCACAATTTGC 27489
Qy 677 CCCAATATCTCCGGGTTTCTCTGAGCTCCATTCCTTGATATCTCAGCTCTTTCAAGT 736
Db 27490 CCCAATATCTCCGGGTTTCTCTGAGCTCCATTCCTTGATATCTCAGCTCTTTCAAGT 27549
Qy 737 ATCTCTCGAGACCGACCATCCATAATTCATTTTGAAGGCCCTTTTAGAGAATCT 796
Db 27550 ATCTCTCGAGACCGACCATCCATAATTCATTTTGAAGGCCCTTTTAGAGAATCT 27609
Qy 797 TATTCCTCAATATTTGACTCTCTGAGTCAATTCAGGAAG 834
Db 27610 TATTCCTCAATATTTGACTCTCTGAGTCAATTCAGGAAG 27647

RESULT 9
US-10-243-735-3
; Sequence 3, Application US/10243735
; Patent No. 6706510
; GENERAL INFORMATION:
; APPLICANT: WEBSTER, Marion et al
; TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC
; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES
; FILE REFERENCE: CL001212-DIV
; CURRENT APPLICATION NUMBER: US/10/243,735
; CURRENT FILING DATE: 2002-09-16
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 63588
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(63588)
; OTHER INFORMATION: n = A,T,C or G
US-10-243-735-3

Query Match 8.7%; Score 205.2; DB 4; Length 63588;
Best Local Similarity 96.3%; Pred. No. 1.2e-48;
Matches 210; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

Qy 617 TCCTTTTGAGGGTAACTTACAGCAGCTGGTTCTGAAAGATTGTCAAGCACAATTTTGC 676
Db 27430 TCAGTTTGAGGGTAACTTACAGCAGCTGGTTCTGAAAGATTGTCAAGCACAATTTTGC 27489
Qy 677 CCCAATATCTCCGGGTTTCTCTGAGCTCCATTCCTTGATATCTCAGCTCTTTCAAGT 736
Db 27490 CCCAATATCTCCGGGTTTCTCTGAGCTCCATTCCTTGATATCTCAGCTCTTTCAAGT 27549
Qy 737 ATCTCTCGAGACCGACCATCCATAATTCATTTTGAAGGCCCTTTTAGAGAATCT 796
Db 27550 ATCTCTCGAGACCGACCATCCATAATTCATTTTGAAGGCCCTTTTAGAGAATCT 27609
Qy 797 TATTCCTCAATATTTGACTCTCTGAGTCAATTCAGGAAG 834
Db 27610 TATTCCTCAATATTTGACTCTCTGAGTCAATTCAGGAAG 27647

Qy 12 ATGGATAAGTACGATGTGATTAAGGCCATCGGCAAGGTCCCTTCGGGAAGCATACTTA 71
Db 1 ATGGAGAAGTACGAGCGGATCCGAGTGTGGGAGAGGTCCCTTCGGGATTTGTCACTG 60
Qy 72 GCTAAAGGGAATTCAGATAGCAAGCACTGTGTCTATAAAGAGATCAATTTTGAAGAATG 131

RESULT 10
US-09-992-481-1
; Sequence 1, Application US/09992481
; Patent No. 6593125
; GENERAL INFORMATION:
; APPLICANT: Friddle, Carl Johan
; APPLICANT: Hilbun, Erin
; APPLICANT: Mathur, Brian
; APPLICANT: Turner, C. Alexander Jr.
; TITLE OF INVENTION: No. 6593125el Human Kinases and Polynucleotides
; FILE REFERENCE: LEX-0266-USA
; CURRENT APPLICATION NUMBER: US/09/992,481
; CURRENT FILING DATE: 2001-11-19
; PRIOR APPLICATION NUMBER: US 60/252,011
; PRIOR FILING DATE: 2000-11-20
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 2079
; TYPE: DNA
; ORGANISM: homo sapiens
US-09-992-481-1

Query Match 7.2%; Score 170.2; DB 4; Length 2079;
Best Local Similarity 52.1%; Pred. No. 2.4e-39;
Matches 404; Conservative 0; Mismatches 368; Indels 3; Gaps 1;

Qy 12 ATGGATAAGTACGATGTGATTAAGGCCATCGGCAAGGTCCCTTCGGGAAGCATACTTA 71
Db 1 ATGGAGAAGTACGAGCGGATCCGAGTGTGGGAGAGGTCCCTTCGGGATTTGTCACTG 60
Qy 72 GCTAAAGGGAATTCAGATAGCAAGCACTGTGTCTATAAAGAGATCAATTTTGAAGAATG 131

Db 61 TGCCTGGAAAGGCTGACCAAGAGCTGGTGTATCATCAAGCAGATTCAGTGGAAACAGATG 120
Qy 132 CCCATACAGAAAAGAGCTTCAAAAGAAAGAGTATCTTCTGGAAAAGATGAACAT 191
Db 121 ACCAAGGAAGAGCGGACGAGCCAGAAATGAGTCCAGAGTCTCAAGCTGCTCAACAC 180
Qy 192 CCCACATGTAGCTTCTCAATTCATTTCAAGAGATGCCAGCTGTTTATTTGTAATG 251
Db 181 CCCATGTCAATGAGTACTAGAGAACTTCTGGAAGACAAAGCCCTTATGATGCCATG 240
Qy 252 GAATATGTGATGAGGGAGTCTCATGAAAAGATCAATAGACAAACGGGTGTTTATTT 311
Db 241 GAATATGCAACGAGCGGACCTCTGGCTGAGTTTCATCCAAAAGCGCTGTAATTCCTGCTG 300
Qy 312 AGTGAAGATCAGATCCTCGGTGGTGTGTTGTACAGATTTCTTAGACATAAACAATATCAT 371
Db 301 GAGGAGGAGACCATCTCGACTCTTCTGTGAGATCCTGCTTGCACTGCAATCATGTGCAC 360
Qy 372 GACAGGAAGATATTACAGAGGACATAAAAGCTCAGAACATTTTCTTAGCAAGAACGGA 431
Db 361 ACCACCTCATCTCCGACCGAGACCTCAAGACCCAGAACATCTCTGCTTGACAAACACCGC 420
Qy 432 ATGGTGGCAAGCTTGGGAGCTTGGTATAGCAAGAGTCTGTAATATTCATGTAATTCATGTA 491
Db 421 ATGGTCTGCAAGATCGGTGATTTCCGGCATCTCCAGATCTCT---TAGCAGCAAGAGCAAG 477
Qy 492 GCTCAAGCTGTATTTGGAACACCTTACTACCTGCTCCCGAGAGATCTGTGAGAAATAAACCC 551
Db 478 GCCTACAGGTGGTGGTACCCCATGCTATATCTCCCTGAGCTGTGTGAGGGCAAGCC 537
Qy 552 TACAAACAATAAAGCATTTTGTGTCTTGGCTGTGTCTTATATGAGCTCTGCACACTT 611
Db 538 TACAAACAGAGAGTACATCTGGGCCCTGGGCTGTCTCTACGAGCTGGCCAGGCTC 597
Qy 612 AAACATCTTTTGAGGTAAACAATTCAGCAGCTGGTCTGGAAGATTTGTCAAGCAGAT 671
Db 598 AAGAGGGCTTCGAGGCTGCGAATTCAGCAGCTGGTCTGGAAGATTCATGAGTGGCACC 657
Qy 672 TTTGCCCAATATCTCCGGGGTTTTCTCGTGAAGCTCCATTCCTTTGATATCTCAGCTCTT 731
Db 658 TTTGACATCTCTGACCGGTACAGCCCTGAGCTTGGCCAGCTGGTCTGAGTCTACTC 717
Qy 732 CAGTATCTCTCGAGACGACCATCCATAATTCATTTTGAAGAGCCCTTTT 786
Db 718 AGCCTGGAGCTGCCCGGCGCCACCTACTGACACATCATGCGACAGCCCTCT 772

RESULT 11

US-08-870-529-1
; Sequence 1, Application US/08870529
; Patent No. 6080557
; GENERAL INFORMATION:
; APPLICANT: Sims, John E.
; APPLICANT: Virca, G. Duke
; APPLICANT: Bird, Timothy A.
; APPLICANT: Anderson, Dirk M.
; TITLE OF INVENTION: IL-1/TNF- α -ACTIVATED KINASE (ITAK),
; METHOD OF INVENTION: AND METHODS OF MAKING AND USING THE SAME
; NUMBER OF SEQUENCES: 18
; CORRESPONDENCE ADDRESS:
; ADDRESS: SEED AND BERRY LLP
; CITY: Seattle
; STATE: Washington
; COUNTRY: USA
; ZIP: 98104
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: IBM PC compatible
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/870,529
; FILING DATE: 06-JUN-1997
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Mcmasters, David D.
; REGISTRATION NUMBER: 33,963
; REFERENCE/DOCKET NUMBER: 480052.418
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (206) 622-4900
; TELEFAX: (206) 682-6031
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2940 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1..2937
; US-08-870-529-1

Query Match 7.0%; Score 166.8; DB 3; Length 2940;
Best Local Similarity 54.8%; Pred. No. 3e-38;
Matches 352; Conservative 0; Mismatches 287; Indels 3; Gaps 1;

Qy 36 GCATCGGGCAAGTGCCTTCGGGAAGCATACATTAGCTAAAGGGAATCAGATAGCAAG 95
Db 169 GTCTGGGCGCGGCGCTTCGGGAAGCGCCAGCTGTACCCGCGCACCGAGATGACTCA 228
Qy 96 CACTGTGTCAAAAAGAGATCAATTTTGAAGAAGATGCCCATACAAGAAAAAGAGTTCA 155
Db 229 CTGTTGTGTGAAGAGAGTGCATTTGACCCGGCTGTCTGAGAAGGAACGTCGTGATGCC 288
Qy 156 AAGAAGAGAGTATTTCTTGTGAAAAGATGAACATCCACATTTGAGCTTCTTCAAT 215
Db 289 TTGAATGAGATAGTTATTTCTGGCACTGCTGCAGCAGCAACATTTATTCCTACTACAT 348
Qy 216 TCATTTCAAGAGAAGTGGCAGGCTGTTTATTTGTAAGTAATTTGTATGGAGGGGATCTC 275
Db 349 CACTTCATGGAACATACACGCTGCTGATGAGCTGGATATTTGTAAGGAGGAACCTG 408
Qy 276 ATGAAAAGAGATCAATAGACAACCGGGTGTGTTATTTAGTGAAGATCAGATCCTCGGTTGG 335
Db 409 TATGACAAAATCCTTCGTCAGAAAGCAAGTTGTTTGAGGAAGAGATGCTGGTGTGTAC 468
Qy 336 TTGTACAGATTTCTTAGGACTAAACATATTCATGACAGGAAGATATTACACAGGAC 395
Db 469 CTATTTCAAGATTTTCAGCAGTGGCTGCATCCATAAAGCTGGATTCCTTCATAGAT 528
Qy 396 ATAAAAGCTCAGAACATTTTCTTAGCAAGAACCGAATGGTGGCAAAAGCTTGGGACTTT 455
Db 529 ATAAAGACATTAATATTTTCTGACCAAG---GCAACCTGATAAAACCTTGAGATTAT 585
Qy 456 GGTATAGGAAGAGTCTGTAATTTCCATGGAACCTTCGTAAGCTTGTATTCGACACT 515
Db 586 GGCCTAGCAAGAAACTTAATTTCTGAGATTCCTAGGCTGAGACGCTTGTGGAAACCCA 645
Qy 516 TACTACTGTCCCGAGAGATCTGTGAGATAAACCCTACAAATATAAAGCGATATTTGG 575
Db 646 TATTACATGCTCCAGAGCTCTGTCAAGGAGTAAAGTACAAATTTCAAGTCTGATATCGG 705
Qy 576 TCTTTGGTGTGTTTATGAGCTCTGCACACTTAACATCTTTTGAGGGTAAACAC 635
Db 706 GCAGTTGGTGGTCAATTTTGAAGCTGTACTTTAAAGAGAGACCTTTGATCTCAAAAC 765
Qy 636 TTACAGCAGCTGTTCTGAGAGATTTGTCAAGCAGATTTTGCC 677
Db 766 CCATTAACCTGTGTGTGAAGATCTGGAAGGAATTCGGGCC 807

RESULT 12

US-09-544-794-1
; Sequence 1, Application US/09544794

```
; Patent No. 6541232
; GENERAL INFORMATION:
; APPLICANT: Sims, John E.
; APPLICANT: Virca, G. Duke
; APPLICANT: Bird, Timothy A.
; APPLICANT: Anderson, Dirk M.
; FILE REFERENCE: 2005-B
; TITLE OF INVENTION: Polypeptides Having Kinase Activity
; CURRENT FILING DATE: 2000-04-07
; PRIOR APPLICATION NUMBER: US/09/544,794
; PRIOR FILING DATE: 1997-06-06
; PRIOR APPLICATION NUMBER: 60/059,979
; PRIOR FILING DATE: 1996-06-10
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1
; LENGTH: 2940
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)...(2940)
US-09-544-794-1

Query Match 7.0%; Score 166.8; DB 4; Length 2940;
Best Local Similarity 54.8%; Pred. No. 3e-38;
Matches 352; Conservative 0; Mismatches 287; Indels 3; Gaps 1;

QY 36 GCATCGGCGAGGTCCTTCGGGAAGCATCTAGCTAAAGGAAATCAGATAGCAAG 95
Db 169 GTCTCGGCGCGGCCCTTCGGGAAGCCACGCTGTACCGCGCACCGAGGATCACTCA 228
QY 96 CACTGTGTCAATAAAGAGATCAATTTTGAAGAAGATGCCATACAAAGAAAAGAACTTCA 155
Db 229 CTGGTTGTGTGAAGGAAGTCGATTGTACCGCGCTGTCTGGAAGGAAGCTGTATGCC 288
QY 156 AGAAGAAGTACTGTTCTTCTGGAAGAAGATGAACATCCCAATTTGTAGCTTTTCAAT 215
Db 289 TTGAATGAGATAGTATTCTGCACTGTCTGCAAGCAAGATTTTGAAGGAAGATGCTGTGTGTAC 348
QY 216 TCATTTCAAGAGAATGGCAGGCTGTTTATTGTAATGGAATATTGTATGAGGAGGATCTC 275
Db 349 CACTTCATGCAATACCAACGCTGTCTGATGAGCTGGATATTGTAAGGAGGAACTG 408
QY 276 ATGAAAAGGATCAATAGACAACGGGGTGTATTATTAGTAGAGATCAGATCCTCGTTGG 335
Db 409 TATGACAAAATCCTTCGTGAGAGGACAAAGTTGTTTGAAGGAAGATGCTGTGTGTGTAC 468
QY 336 TTTGTACAGATTTCTTAGGACTTAAACATATTTCATGACAGGAATATTACAGGGAC 395
Db 469 CTATTTCAAGATTTTTCAGAGTGTGCTGATCCATTAAGCTGGAATCCTTCATAGAT 528
QY 396 ATAAAAGCTCAGAACATTTTCTTAGCAAGACGGAATGGTGGCAAAAGCTTGGGACATT 455
Db 529 ATAAAAGACATTAATATTTTCTGACCAAG--GCAAACTGATAAACTTTGGAGATTAT 585
QY 456 GGTATGCAAGAGTCTGTAATTTCCATGGAATGCTGCAACTGCTGCACTGATTTGGAACCT 515
Db 586 GGCCTAGCAAGAACTTAATTTCTGAGTATTCATGCTGAGACGCTTGTGGGAACCCCA 645
QY 516 TACTACCTGTCCCGCAGAGATCTGTGAGAAATAAACCTTACAACATATAAAGCGATTTGG 575
Db 646 TATTACATGCTCTCCAGAGCTCTGTGACAGGAGTAAAGTACATTTCAAGCTGTATCTGG 705
QY 576 TCTCTTGGCTGTCTTATAGCTCTGCACTTAAACATTCCTTTTGGAGGGTAAACAAC 635
Db 706 GCAATGCTGCTGCTCAATTTTGAAGTCTTACCTTAAAGAGAGAGCTTTGATGCTACAAAC 765
QY 636 TTACAGCAGCTGGTTCTGAAGATTTGTCAAGACATTTTGGC 677
Db 766 CCACCTAACCTGTGTGTGAAGATCTGTGCAAGGAATTCGGGCC 807

RESULT 13
US-08-870-529-8
; Sequence 8, Application US/08870529
; Patent No. 6080557
; GENERAL INFORMATION:
; APPLICANT: Sims, John E.
; APPLICANT: Virca, G. Duke
; APPLICANT: Bird, Timothy A.
; APPLICANT: Anderson, Dirk M.
; TITLE OF INVENTION: IL-1/TNF- $\alpha$ -ACTIVATED KINASE (ITAK),
; NUMBER OF SEQUENCES: 18
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SEED and BERRY LLP
; STREET: 6300 Columbia Center, 701 Fifth Avenue
; CITY: Seattle
; STATE: Washington
; COUNTRY: USA
; ZIP: 98104
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/870,529
; FILING DATE: 06-JUN-1997
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: McMasters, David D.
; REGISTRATION NUMBER: 33,963
; REFERENCE/DOCKET NUMBER: 480052.418
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (206) 622-4900
; TELEFAX: (206) 682-6031
; INFORMATION FOR SEQ ID NO: 8:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3264 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
US-08-870-529-8

Query Match 7.0%; Score 166.8; DB 3; Length 3264;
Best Local Similarity 54.8%; Pred. No. 3.3e-38;
Matches 352; Conservative 0; Mismatches 287; Indels 3; Gaps 1;

QY 36 GCATCGGCGAGGTCCTTCGGGAAGCATCTAGCTAAAGGAAATCAGATAGCAAG 95
Db 288 GTCTCGGCGCGGCCCTTCGGGAAGCCACGCTGTACCGCGCACCGAGGATCACTCA 347
QY 96 CACTGTGTCAATAAAGAGATCAATTTTGAAGAAGATGCCATACAAAGAAAAGAACTTCA 155
Db 348 CTGGTTGTGTGGAAGGAGTCTGATTGTGACCGCGCTGTCTGAGAGGAACGTCGTGTATGCC 407
QY 156 AGAAGAAGTACTTCTTCTGGAAGAAGATGAACATCCCAACATTTGTAGCTTTTCAAT 215
Db 408 TTGAATGAGATAGTATTCTTGGCACTGTCTGCAAGCAACATTTATTCCTACTACAAT 467
QY 216 TCATTTCAAGAGAATGGCAGGCTGTTTATTGTAATGGAATATTGTATGAGGAGGATCTC 275
Db 468 CACTTCATGCAATACCAACGCTGTCTGATGAGCTGGATATTGTAATGAGGGAACCTG 527
QY 276 ATGAAAAGGATCAATAGACAACGGGGTGTATTATTAGTAGAGATCAGATCCTCGTTGG 335
Db 528 TATGACAAAATCCTTCGTGAGAGGACAAAGTTGTTTGAAGGAAGATGCTGTGTGTGTAC 587
QY 336 TTTGTACAGATTTCTTAGGACTTAAACATATTTCATGACAGGAGATATTACAGGGAC 395
Db 588 CTATTTCAAGATTTTTCAGCAGTGTGCTGATCCATTAAGCTGGAATCCTTCATAGAT 647
QY 396 ATAAAAGCTCAGAACATTTTCTTAGCAAGAACGGAATGGTGGCAAAAGCTTGGGGACATT 455
```

Db 648 ATAAAGACATTAATAATTTTCTGACCAAG--GCAACCTGATAAACTTGGAGATTAT 704
Qy 456 GGTATAGCAAGAGTCCCTGAATAATTCATGGAACCTTGTCTCGAATCTTGTATTGGAAACACCT 515
Db 705 GGCCTAGCAAGAAACTTAATCTCGAGTATTCATGAGTCTGAGACGCTTGTGGGAACCCCA 764
Qy 516 TACTACCTGCTCCAGAGATCTGTGAGATAAACCCTAGCAACATAAAACGGATATTGG 575
Db 765 TATTACATGCTCCAGAGCTCTGTCAAGAGTAAAGTACAAATTCAGGCTGATATCTGG 824
Qy 576 TCTCTTGGCTGTCTTATATAGCTCTGCACACTTAAACATCCTTTTGAAGGTAACAAC 635
Db 825 GCAGTTGGCTGCTCATTTTGAACCTTACCTTAAAGAGGACGTTGATGCTACAAAC 884
Qy 636 TTACAGAGCTGTCTTGAAGATTGTCAGACACATTTTGGC 677
Db 885 CCACTTAACCTGTGTGCAAGATCGTCAAGGAATTCGGGCC 926

RESULT 14
US-09-544-794-8
; Sequence 8, Application US/09544794
; Patent No. 6541232
; GENERAL INFORMATION:
; APPLICANT: Sims, John E.
; APPLICANT: Virca, G. Duke
; APPLICANT: Bird, Timothy A.
; APPLICANT: Anderson, Dirk M.
; TITLE OF INVENTION: Polypeptides Having Kinase Activity
; FILE REFERENCE: 2005-B
; CURRENT APPLICATION NUMBER: US/09/544,794
; CURRENT FILING DATE: 2000-04-07
; PRIOR APPLICATION NUMBER: 08/870,529
; PRIOR FILING DATE: 1997-06-06
; PRIOR APPLICATION NUMBER: 60/059,979
; PRIOR FILING DATE: 1996-06-10
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 8
; LENGTH: 3264
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-544-794-8

Query Match 7.0%; Score 166.8; DB 4; Length 3264;
Best Local Similarity 54.8%; Pred. No. 3.3e-38;
Matches 352; Conservative 0; Mismatches 287; Indels 3; Gaps 1;
Qy 36 GCCATCGGGGCAAGTGCCTTCGGGAAGCATACTTAGCTAAGGGAATCAGATAGCAAG 95
Db 288 GTCCTGGCCGCGCGCTTCGGGGAAGCCACGCTGTACCGCCGCAACCGAGGATGACTCA 347
Qy 96 CACTGTGTCTATAAAGAGATCAATTTTGAAGAAGTGCCTATACAGAAAGAAAGCTTCA 155
Db 348 CTGGTGTGTGAGGAGTGTGATTTGACCGGCTGTCTGAGAGGAACTCGTGTATGCC 407
Qy 156 AAGAAAGAGTGTCTTCTCGAAAGATGAACATCCCAACATTTAGTCCCTTCTCAAT 215
Db 408 TTGAATGAGATAGTATTCTGGCACTGTGCGAGCAGCAACATATTGCTACTACAAT 467
Qy 216 TCATTTCAAGAGATGGCAGCTCTTATTCTGAATGGAATTTGTGATGAGGGGATCTC 275
Db 468 CACTTCATGGCAATACACCGCTCTGATGAGCTGGAAATTTGTATGAGGGAACCTG 527
Qy 276 ATGAAAGGATCAATFAGACACCGGGTGTGTATTATTAGTGAAGATCAGATCCTCGGTGG 335
Db 528 TATGACAAAATCCCTTCGTGAGAGGACAAAGTTGTTGAGGAAGAGATGGTGTGTGCTAC 587
Qy 336 TTTGTACAGATTCTCTAGGACTAAACATATTATCATGACGGAAGATATTACACAGGAC 395
Db 588 CTATTTACAGATTGTTTCAGCAGTGTGCTGCATTCATTAAGCTGGAATCTTTCATAGAT 647

Qy 396 ATAAAGCTCAGAACATTTTCTTAGCAAGAACGGAATGTGCAAGCTTGGGACCTTT 455
Db 648 ATAAAGACATTAATAATTTTCTGACCAAG---GCAACCTGATAAACTTGGAGATTAT 704
Qy 456 GGTATAGCAAGAGTCCCTGAATAATTCATGGAACCTTGTCTCGAATCTTGTATTGGAAACCT 515
Db 705 GGCCTAGCAAGAAACTTAATCTCGAGTATTCATGAGTCTGAGACGCTTGTGGGAACCCCA 764
Qy 516 TACTACCTGCTCCAGAGATCTGTGAGATAAACCCTAGCAACATAAAACGGATATTGG 575
Db 765 TATTACATGCTCCAGAGCTCTGTCAAGAGTAAAGTACAAATTCAGGCTGATATCTGG 824
Qy 576 TCTCTTGGCTGTCTTATATAGCTCTGCACACTTAAACATCCTTTTGAAGGTAACAAC 635
Db 825 GCAGTTGGCTGCTCATTTTGAACCTTACCTTAAAGAGGACGTTTGTGCTACAAAC 884
Qy 636 TTACAGAGCTGTCTTGAAGATTGTCAGACACATTTTGGC 677
Db 885 CCACTTAACCTGTGTGCAAGATCGTCAAGGAATTCGGGCC 926

RESULT 15
US-10-196-927-3
; Sequence 3, Application US/10196927
; Patent No. 6797510
; GENERAL INFORMATION:
; APPLICANT: Walke, D. Wade
; APPLICANT: Scoville, John
; APPLICANT: Friddle, Carl Johan
; TITLE OF INVENTION: No. 6797510el Human Kinases and Polynucleotides Encoding the Same
; FILE REFERENCE: LEX-0348-USA
; CURRENT APPLICATION NUMBER: US/10/196,927
; CURRENT FILING DATE: 2002-05-20
; PRIOR APPLICATION NUMBER: US 60/293,248
; PRIOR FILING DATE: 2001-05-24
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 1449
; TYPE: DNA
; ORGANISM: homo sapiens
US-10-196-927-3

Query Match 5.6%; Score 131.6; DB 4; Length 1449;
Best Local Similarity 51.3%; Pred. No. 4.1e-28;
Matches 364; Conservative 0; Mismatches 334; Indels 12; Gaps 2;
Qy 102 GTCATAAAAGAGATCAATTTTGAAGAAGTGCCTATACAGAAAGAAAGAGCTTCAAAGAAA 161
Db 175 GTACTTAAGGAATATCTGTTGGAGAACTAAATCCAAATGAAGCTGTACAGGCCAATTG 234
Qy 162 GAAGTGATCTCTCGAAAGATGAACATCCCAACATTTAGCCTTCTTCAATTCATTT 221
Db 235 GAAGCCCAACTCTCTCCCAAGCTGGACCCACCCAGCATTTGCAAGTTCCATGCAAGTTT 294
Qy 222 CAAGAGATGGCAGCTGTATTCTTAATGGAATATTGTGATGGAGGGATCTCATGAAA 281
Db 295 GTGAGCAAGATAATTTCTGCATTATCAGGAGTACTGTGAGGGCCGAGATCTGGACGAT 354
Qy 282 AGGATCA-----ATAGACAACGGGTGTGTTATTAGTGAAGATCAGATCCTCGGTGG 335
Db 355 AAAATTCAAGGAATATAAAACAAGCTGGAATAATCTTTCCAGAAATCAAAATATAGAAATGG 414
Qy 336 TTTGTACAGATTCTCTAGGACTAAACATATTCTATGACAGGAAGATATTACAGGGAC 395
Db 415 TTATTCAGCTCTCTCTGGAGTCTGCTACATGATGATGAGAGGAGATATTCAATCGAGAC 474
Qy 396 ATAAAGCTCAGAACATTTTCTTAGCAAGAACCGGAATGGTGGCAAGCTTGGGACTTT 455
Db 475 TTAAGCTCAAGAAATGTATTTCTGAAAAATAA-----TCTCCTTAAATTTGGAGATT 528
Qy 456 GGTATAGCAAGAGTCTCTGATTAATTCATGGAATCTGCTGGAACCTTGTATTGGAAACCT 515


```

Db      529 GGAGTTTCTCGATTCTTAATGGGATCCTGTGACCTGGCCACAACTTTAACTGGAACTCCC 588
Qy      516 TACTACCTGTCCCGAGAGATCTGTCAAGATAAACCCCTACACAAATAAAACGGATATTTGG 575
Db      589 CATTATATGAGTCTCTGAGGCTCTGAAACACCAAGGCTATGACACAAAGTCGGACATCTGG 648
Qy      576 TCTCTTGGCTGTGTCTTATATGAGCTCTGCACACTTAAACATCCTTTTGAGGGTAACAAC 635
Db      649 TCACTGGCATGCAATTTTGTATGAGATGTGCTGCATGAATCATGCAATTCGCTGGCTCCAAT 708
Qy      636 TTACAGCAGCTGGTTCTGAAGATTGTCAAGCACATTTTGCCCAATATCTCCGGGGTTT 695
Db      709 TTCTTATCCATTGTTTAAATAATTGTTGAAGGTGACACACCTTCTCTCCCTGAGAGATAT 768
Qy      696 TCTCGTGAGCTCCATTCCCTTGATATCTGAGCTCTTTCAAGTATCTCCTCGAGACCGACCA 755
Db      769 CCAAAAGAACTAAATGCCATCATGGAAGCATGTTGAACAGAAATCCTTCATTAAAGACCA 828
Qy      756 TCCATAAATTCATTTTGAAGGCCCTTTTATAGAGATCTTTATCCCAA 805
Db      829 TCTGCTATCGAAATTTTAAATAATCCCTTACCTTGATGAGCAGCTACAGAA 878

```

Search completed: November 8, 2004, 02:03:08
 Job time : 222 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.
OM nucleic - nucleic search, using sw model
Run on: November 8, 2004, 00:08:24 ; Search time 4618 Seconds
(without alignments)
2763.178 Million cell updates/sec

Title: US-10-730-010-1
Perfect score: 2370
Sequence: 1 cattggagaccatgataag.....ctgtcatcgcttaacgacgc 2370

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 3611042 seqs, 2692057975 residues
Total number of hits satisfying chosen parameters: 7222084

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications NA.*
1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq.*
2: /cgn2_6/ptodata/1/pubpna/PCT_NEW_PUB.seq.*
3: /cgn2_6/ptodata/1/pubpna/US06_NEW_PUB.seq.*
4: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq.*
5: /cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq.*
6: /cgn2_6/ptodata/1/pubpna/PCTUS_PUBCOMB.seq.*
7: /cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq.*
8: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq.*
9: /cgn2_6/ptodata/1/pubpna/US09A_PUBCOMB.seq.*
10: /cgn2_6/ptodata/1/pubpna/US09B_PUBCOMB.seq.*
11: /cgn2_6/ptodata/1/pubpna/US09C_PUBCOMB.seq.*
12: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq.*
13: /cgn2_6/ptodata/1/pubpna/US10A_PUBCOMB.seq.*
14: /cgn2_6/ptodata/1/pubpna/US10B_PUBCOMB.seq.*
15: /cgn2_6/ptodata/1/pubpna/US10C_PUBCOMB.seq.*
16: /cgn2_6/ptodata/1/pubpna/US10D_PUBCOMB.seq.*
17: /cgn2_6/ptodata/1/pubpna/US10E_PUBCOMB.seq.*
18: /cgn2_6/ptodata/1/pubpna/US10_NEW_PUB.seq.*
19: /cgn2_6/ptodata/1/pubpna/US11_NEW_PUB.seq.*
20: /cgn2_6/ptodata/1/pubpna/US60_NEW_PUB.seq.*
21: /cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2370	100.0	2370	14	US-10-243-735-1
2	2370	100.0	2370	17	US-10-730-010-1
3	1665	70.3	2240	9	US-09-940-921B-5
4	1865	70.3	2816	18	US-10-618-941-28
5	1656.2	69.9	1965	9	US-09-940-921B-3
6	1654	69.8	2052	9	US-09-940-921B-1
7	1199.6	50.6	1587	16	US-10-114-270-187
8	1176.8	49.7	1453	16	US-10-114-270-185
9	369.4	15.6	5426	9	US-09-783-320-50
10	369.4	15.6	5448	15	US-10-037-270-246
11	369.4	15.6	5448	15	US-10-117-722-246
12	369.4	15.6	5532	15	US-10-037-270-245

13	369.4	15.6	5532	15	US-10-117-722-245
14	369.4	15.6	5583	18	US-10-618-941-29
15	368	15.5	3645	9	US-09-783-320-3
16	368	15.5	3645	18	US-10-757-262-123
17	262.2	11.1	1588	16	US-10-114-270-183
18	260.6	11.0	1781	16	US-10-114-270-181
19	260.6	11.0	2257	16	US-10-114-270-179
20	254.2	10.7	2104	18	US-10-757-262-117
21	254.2	10.7	2110	18	US-10-415-011-29
22	254.2	10.7	2326	18	US-10-618-941-30
23	238.4	10.1	1846	9	US-09-870-962-15
24	205.2	8.7	63588	14	US-10-243-735-3
25	205.2	8.7	63588	17	US-10-730-010-3
26	170.2	7.2	840	14	US-10-162-706-13
27	170.2	7.2	1383	14	US-10-162-706-3
28	170.2	7.2	1328	15	US-10-288-798-47
29	170.2	7.2	1528	16	US-10-362-892-47
30	170.2	7.2	2079	9	US-09-992-481-1
31	170.2	7.2	2079	15	US-10-434-034-1
32	170.2	7.2	2142	14	US-10-162-706-1
33	168.6	7.1	1037	16	US-10-074-978A-15
34	168.6	7.1	2079	9	US-09-910-150-6
35	168.6	7.1	2079	16	US-10-377-097-6
36	168.6	7.1	2438	9	US-09-910-150-4
37	168.6	7.1	2438	16	US-10-377-097-4
38	168.4	7.1	2940	9	US-09-884-001-3
39	167.4	7.1	3696	17	US-10-620-052A-31
40	166.8	7.0	2940	15	US-10-372-736-1
41	166.8	7.0	3264	15	US-10-372-736-8
42	165.8	7.0	3544	16	US-10-425-114-26260
43	165.8	7.0	3544	16	US-10-425-114-26275
44	165.8	7.0	3698	15	US-10-172-118-880
45	165.8	7.0	3698	16	US-10-342-887-880

ALIGNMENTS

RESULT 1

US-10-243-735-1
; Sequence 1, Application US/10243735
; Publication No. US20030022341A1
; GENERAL INFORMATION:
; APPLICANT: WEBSTER, Marion et al
; TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC
; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES
; TITLE OF INVENTION: THEREOF
; FILE REFERENCE: CL001212DIV
; CURRENT APPLICATION NUMBER: US/10/243,735
; CURRENT FILING DATE: 2002-09-16
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 2370
; TYPE: DNA
; ORGANISM: Human
US-10-243-735-1

Query Match	100.0%;	Score 2370;	DB 14;	Length 2370;	
Best Local Similarity	100.0%;	Pred. No. 0;			
Matches 2370;	Conservative	0;	Mismatches	0;	
		Indels	0;	Gaps	0;
Qy	1	CATTGAGACCATGGATAAGTACGATGTGATTAAAGCCATCGGCGAAGGTGCTTCGGGA	60		
Dd	1	CATTGAGACCATGGATAAGTACGATGTGATTAAAGCCATCGGCGAAGGTGCTTCGGGA	60		
Qy	61	AAGCATCTTAGCTAAAGGGAATCAGATAGCAAGCACTGTGTCTATAAAGAGATCAATT	120		
Dd	61	AAGCATCTTAGCTAAAGGGAATCAGATAGCAAGCACTGTGTCTATAAAGAGATCAATT	120		
Qy	121	TTGAAAAGATGCCCATACAGAAAAGAGCTTCAAGAAAAGAGTATCTTCTCGAAA	180		
Dd	121	TTGAAAAGATGCCCATACAGAAAAGAGCTTCAAGAAAAGAGTATCTTCTCGAAA	180		

1141 ATCACCTTCTCCTCAGAAATACCTGGATTGAGGATTACGGTCAGGAAACGAGGCATG 1200
Db ATCACCTTCTCCTCAGAAATACCTGGATTGAGGATTACGGTCAGGAAACGAGGCATG 1293
Qy GTCCATCCCAAGTCAATGSCCTGCTGAGTACCTTCAGAGAAAATTTGAAGCTCAACAAT 1260
Db GTCCATCCCAAGTCAATGSCCTGCTGAGTACCTTCAGAGAAAATTTGAAGCTCAACAAT 1353
Qy ATAAAGTTGAAGTGGAGAACCAATTTGGTCTTCATCTCTGCGAGCCCAATTA 1320
Db ATAAAGTTGAAGTGGAGAACCAATTTGGTCTTCATCTCTGCGAGCCCAATTA 1413
Qy ACCAGAGACAAAGAGCTAAAGAGTAATGGAGAGAGCTAGATTCCAGGAGCTGCCATTTA 1380
Db ACCAGAGACAAAGAGCTAAAGAGTAATGGAGAGAGCTAGATTCCAGGAGCTGCCATTTA 1473
Qy GGAAGAACGAAATGAAGAGAACAGGAATATTTGAAGCAGTTAGAGAAAATACGCCAACAGT 1440
Db GGAAGAACGAAATGAAGAGAACAGGAATATTTGAAGCAGTTAGAGAAAATACGCCAACAGT 1533
Qy ACCTCAATGACATGAAGAGAAATTAGAAGAGAGATGGGAGAGAACCCAG----- 1488
Db ACCAATGACATGAAGAGAAATTAGAAGAGAGATGGGAGAGAACCCAGAGGAGAACTCAA 1593
Qy ----- 1488
Db AAATAAGTCATAAAACCTATTGGTGAAGAGAGTAACCTGCTCCTCAAGATGCAT 1653
Qy -----AGGACATTGAAAGAGACTTGAAACAAATGAGGCTTCAGA 1527
Db CTGAGGAGAGACCTGTGCGAGGACATTGAAAGAGACTTGAAACAAATGAGGCTTCAGA 1713
Qy ACAAGAGGAGAAAGTAAAAATCCAGAACAGAAATATAAGCTAGAGAGGGGGTAAAAATTTG 1587
Db ACAAGAGGAGAAAGTAAAAATCCAGAACAGAAATATAAGCTAGAGAGGGGGTAAAAATTTG 1773
Qy AAATTAATTTAGACAAATGATTTCTGATGAAACATCTCTCAGAGAGAGAGGAGGAAATGG 1647
Db AAATTAATTTAGACAAATGATTTCTGATGAAACATCTCTCAGAGAGAGAGGAGGAAATGG 1833
Qy ATATACCAATGAACTTTGACCTTTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1707
Db ATATACCAATGAACTTTGACCTTTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1893
Qy TAAAGGAGCAGGAGATATACAGACAAAGCATTGAAACCTTCACTCCAGAGAGCAG 1767
Db TAAAGGAGCAGGAGATATACAGACAAAGCATTGAAACCTTCACTCCAGAGAGCAG 1953
Qy CATTTACAGAACTGACT 1784
Db GGTTCACGAGACT 1970

RESULT 4

US-10-618-941-28
; Sequence 28, Application US/10618941
; Publication No. US2004019792A1
; GENERAL INFORMATION:
; APPLICANT: WHYTE, DAVID
; APPLICANT: MANNING, GERARD
; APPLICANT: CAENEPEEL, SEAN
; TITLE OF INVENTION: NOVEL KINASES
; FILE REFERENCE: 034536-0321
; CURRENT APPLICATION NUMBER: US/10/618,941
; CURRENT FILING DATE: 2003-07-15
; PRIOR APPLICATION NUMBER: 60/395,632
; PRIOR FILING DATE: 2002-07-15
; NUMBER OF SEQ ID NOS: 143
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 28
; LENGTH: 2816
; TYPE: DNA
; ORGANISM: Homo sapiens

US-10-618-941-28
Query Match 70.3%; Score 1665; DB 18; Length 2816;
Best Local Similarity 94.5%; Pred. No. 0;
Matches 1774; Conservative 0; Mismatches 10; Indels 93; Gaps 1;
Qy 1 CATTGGAGACCATGGATAAGTACGATGTGATTAAAGGCCATCGGGCAAGGTGCCTTCGGGA 60
Db 307 CATTGGAGACCATGGATAAGTACGATGTGATTAAAGGCCATCGGGCAAGGTGCCTTCGGGA 366
Qy 61 AAGCATCTTACGTAAAGGGAAATCAGATAGCAAGCACTGTGTCTATAAAGAGATCAATT 120
Db 367 AAGCATCTTACGTAAAGGGAAATCAGATAGCAAGCACTGTGTCTATAAAGAGATCAATT 426
Qy 121 TTGAAAAGATGCCCATACAGAAAAGAGCTTCARAGAAAGAGTCAATCTCTCGAAA 180
Db 427 TTGAAAAGATGCCCATACAGAAAAGAGCTTCARAGAAAGAGTCAATCTCTCGAAA 486
Qy 181 AGATGAAACATCCCAACATTTAGCCTTCTTCAATTTCAATTTCAAGAGAAATGGCAGGTGT 240
Db 487 AGATGAAACATCCCAACATTTAGCCTTCTTCAATTTCAATTTCAAGAGAAATGGCAGGTGT 546
Qy 241 TTATTGTAATCGAATATTGTGATGGAGGGATCTCATGAAAAGGATCAATAGACAACGGG 300
Db 547 TTATTGTAATCGAATATTGTGATGGAGGGATCTCATGAAAAGGATCAATAGACAACGGG 606
Qy 301 GTGTGTTATTTAGTGAAGATCAGATCCTCGGTTGGTTGTACAGATTTCTCTAGGACTAA 360
Db 607 GTGTGTTATTTAGTGAAGATCAGATCCTCGGTTGGTTGTACAGATTTCTCTAGGACTAA 666
Qy 361 AACATTTATCATGACAGAGAAATATTACACAGGACATAAAGCTCAGAACATTTTCTTA 420
Db 667 AACATTTATCATGACAGAGAAATATTACACAGGACATAAAGCTCAGAACATTTTCTTA 726
Qy 421 GCAAGACGGAATGCTGCAAGCTTTGGGAGCTTTGGTATAGCAAGAGTCTGTAATATT 480
Db 727 GCAAGACGGAATGCTGCAAGCTTTGGGAGCTTTGGTATAGCAAGAGTCTGTAATATT 786
Qy 481 CCATGGAATCTGCTCGAACTTTGTAATGGAAACACCTTACTACTCTGCTGCCAGAGATCTGTC 540
Db 787 CCATGGAATCTGCTCGAACTTTGTAATGGAAACACCTTACTACTCTGCTGCCAGAGATCTGTC 846
Qy 541 AGAATACCCCTACACAAATTAACGATTTGGTCTCTGGCTGTGTCTTATATAGC 600
Db 847 AGAATACCCCTACACAAATTAACGATTTGGTCTCTGGCTGTGTCTTATATAGC 906
Qy 601 TCTGCACACTTTAAACATCTTTTGGAGGTAAACAACTTACAGCAGCTGGTCTGAAGATTT 660
Db 907 TCTGCACACTTTAAACATCTTTTGGAGGTAAACAACTTACAGCAGCTGGTCTGAAGATTT 966
Qy 661 GTCAAGCAGATTTGGCCCCAAATATCTCGGGGGTTTTCTCGTGAAGCTCAATTCCTTGATAT 720
Db 967 GTCAAGCAGATTTGGCCCCAAATATCTCGGGGGTTTTCTCGTGAAGCTCAATTCCTTGATAT 1026
Qy 721 CTCAGCTCTTTCAAGTATCTCTCGAGACGACCATCCATAAAATTCATTTTGAAGAGGC 780
Db 1027 CTCAGCTCTTTCAAGTATCTCTCGAGACGACCATCCATAAAATTCATTTTGAAGAGGC 1086
Qy 781 CTTTATTAGAAATCTTTATTTCCCAATATTGACTCTCTGAGGTCAATTCAGAGAAATTC 840
Db 1087 CTTTATTAGAAATCTTTATTTCCCAATATTGACTCTCTGAGGTCAATTCAGAGAAATTC 1146
Qy 841 GTCACATGTTTATATGACAGCAGAGCGCCAGCTTCTCGACATGCTGGGAGGTGCTCC 900
Db 1147 GTCACATGTTTATATGACAGCAGAGCGCCAGCTTCTCGACATGCTGGGAGGTGCTCC 1206
Qy 901 AGAAGTGTAAATACAAAAGTGAATTTCCGGGGAAAGTGCACCACCAAGATCAAGGATAT 960
Db 1207 AGAAGTGTAAATACAAAAGTGAATTTCCAGGGAAAGTGCACCACCAAGATCAAGGATAT 1266
Qy 961 CTGTCCCAATTAAGAGATCTATATTGCAATAGAAATGAATGGAGACCAACCACTGGAG 1020
Db 1267 CTGTCCCAATTAAGAGATCTATATTGCAATAGAAATGAATGGAGACCAACCACTGGAG 1326


```
QY 1021 CCCAGAGCCAGATCTATATAAATATAGTAAGAACCCAAATTTGCTGCTCTGTGGAC 1080
D 1327 CCCAGAGCCAGATCTATATAAATATAGTAAGAACCCAAATTTGCTGCTGTGGAC 1386
QY 1081 ATTATGATTATTATTATGCTCAACTTGTATGCTGAGGAGAGAGCCACAAACCAAGTT 1140
D 1387 ATTATGATTATTATTATGCTCAACTTGTATGCTGAGGAGAGAGCCACAAACCAAGTT 1446
QY 1141 ATCACCCCTATTCTCAAGAAATATCTGGAGTTGAGGATTACGGTCAGAAACGAGGATG 1200
D 1447 ATCACCCCTATTCTCAAGAAATATCTGGAGTTGAGGATTACGGTCAGAAACGAGGATG 1506
QY 1201 GTCATCCCGCAATCAATGGCTGCTGATGATCTTCAAGAAATTTGAAGCTCAACAT 1260
D 1507 GTCCATCCCGCAATCAATGGCTGCTGATGATCTTCAAGAAATTTGAAGCTCAACAT 1566
QY 1261 ATAGTTGAAAGTGAGAGCAATTTGGGTCTTCTGCTCAATCTTCTGCGAGCCAAATTACA 1320
D 1567 ATAGTTGAAAGTGAGAGCAATTTGGGTCTTCTGCTCAATCTTCTGCGAGCCAAATTACA 1626
QY 1321 ACCAGAGACAGAGCTAAGAGTAATGGAGAGAGCCCTAGATTTCAGAGAGCTGCCATTTA 1380
D 1627 ACCAGAGACAGAGCTAAGAGTAATGGAGAGAGCCCTAGATTTCAGAGAGCTGCCATTTA 1686
QY 1381 GGAAGAAACGAATGAAGGAACAGAGTAATTTGGAAGCAGTTAGAGGAATACGCCAACAGT 1440
D 1687 GGAAGAAACGAATGAAGGAACAGAGTAATTTGGAAGCAGTTAGAGGAATACGCCAACAGT 1746
QY 1441 ACCTCAATGACATGAAGAAATATAGAAAGAGATGGGAGAGAACCCAG - 1488
D 1747 ACCAATGACATGAAGAAATATAGAAAGAGATGGGAGAGAACCCAGAGGAGAACTCAA 1806
QY 1489 ----- 1488
D 1807 AATAAGTCAATAAACCTATTTTGGTGAAGAGAGTAACCTGCTGCTCAATCAAGATGCAT 1866
QY 1489 -----AGGACATTGAAAGAGCTTGAAACAAATAGGGCTTCAGA 1527
D 1867 CTGAGGGAGAGCACCTGTGAGGAGACATTGAAAGAGCTTGAAACAAATAGGGCTTCAGA 1926
QY 1528 ACACAAAGGAAGTAATAATCCAGACAGAAATATAGCTTAAGATAAGAGGGGTAAATTTG 1587
D 1927 ACACAAAGGAAGTAATAATCCAGACAGAAATATAGCTTAAGATAAGAGGGGTAAATTTG 1986
QY 1588 AAATTAATTTAGACAAATGTTATTTCTGATGAAACATCTCCAGAGAGAGAGGCAATGG 1647
D 1987 AAATTAATTTAGACAAATGTTATTTCTGATGAAACATCTCCAGAGAGAGAGGCAATGG 2046
QY 1648 ATATACCAATGAACCTTTGACCTTTGAGGATGGATGGAATTTAAGGAATATGATGTTG 1707
D 2047 ATATACCAATGAACCTTTGACCTTTGAGGATGGATGGAATTTAAGGAATATGATGTTG 2106
QY 1708 TAAAGGAGCATGGAGATTATACAGACAAAGCATTTGAAAACTTCACTGCCAGAGACAG 1767
D 2107 TAAAGGAGCATGGAGATTATACAGACAAAGCATTTGAAAACTTCACTGCCAGAGACAG 2166
QY 1768 CATTTACAGACTGACT 1784
D 2167 GGTTCACGAGACT 2183
```

RESULT 5

```
US-09-940-921B-3
; Sequence 3, Application US/09940921B
; Patent No. US20020147320A1
; GENERAL INFORMATION:
; APPLICANT: Fiddle, Carl Johan
; APPLICANT: Hilbun, Erin
; APPLICANT: Nepomichy, Boris
; APPLICANT: Hu, Yi
; TITLE OF INVENTION: No. US20020147320A1el Human Kinase Proteins and Polynucleotides
; FILE REFERENCE: LEX-0227-USA
```

```
; CURRENT APPLICATION NUMBER: US/09/940,921B
; CURRENT FILING DATE: 2002-05-21
; PRIOR APPLICATION NUMBER: US 60/229,280
; PRIOR FILING DATE: 2000-08-31
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 1965
; TYPE: DNA
; ORGANISM: homo sapiens
US-09-940-921B-3
```

```
Query Match 69.9%; Score 1656.2; DB 9; Length 1965;
Best Local Similarity 96.1%; Pred. No. 0;
Matches 1710; Conservative 0; Mismatches 63; Indels 6; Gaps 1;

QY 12 ATGGATAAGTACGATGTGATTAAGGCCATCGGGCAAGTGCCTTCGGGAAAGCATACTTA 71
D 1 ATGGATAAGTACGATGTGATTAAGGCCATCGGGCAAGTGCCTTCGGGAAAGCATACTTA 60
QY 72 GCTAAGGGAATCAGATAGCAAGCACTCTGTGCATATAAAGAGATCAATTTTGAAGATG 131
D 61 GCTAAGGGAATCAGATAGCAAGCACTCTGTGCATATAAAGAGATCAATTTTGAAGATG 120
QY 132 CCATACAAGAAAAAGAACCTTCAAGAAAGAAAGTATTCTTCTGAAAAAGATGAACAT 191
D 121 CCATACAAGAAAAAGAACCTTCAAGAAAGAAAGTATTCTTCTGAAAAAGATGAACAT 180
QY 192 CCCAATTTGTAGCTTCTTCAATTCATTTCAAGAGATGGCAGGCTGTTATTGTAATG 251
D 181 CCCAATTTGTAGCTTCTTCAATTCATTTCAAGAGATGGCAGGCTGTTATTGTAATG 240
QY 252 GAATATTGTGAGGAGGATCTCATGAAAGAGATCAATAGACAAACGGGCTGTGTTATT 311
D 241 GAATATTGTGAGGAGGATCTCATGAAAGAGATCAATAGACAAACGGGCTGTGTTATT 300
QY 312 AGTGAAGATCAGATCTCTGGTTGGTTGTACAGATTCTTAGGACTAATAACATATTAT 371
D 301 AGTGAAGATCAGATCTCTGGTTGGTTGTACAGATTCTTAGGACTAATAACATATTAT 360
QY 372 GACGGAAGATATACACAGGAGACATAAAGCTCAGAACATTTTCTTAGCAAGAACGGA 431
D 361 GACGGAAGATATACACAGGAGACATAAAGCTCAGAACATTTTCTTAGCAAGAACGGA 420
QY 432 ATGTGGCAAAAGCTTTGGGACCTTTGGTATAGCAAGAGTCTTGAATTAATTCATGGA 491
D 421 ATGTGGCAAAAGCTTTGGGACCTTTGGTATAGCAAGAGTCTTGAATTAATTCATGGA 480
QY 492 GCTCGAATTTGTTGGACACCTTACTACCTGTCGCCAGAGATCTGTCAGATTAACCC 551
D 481 GCTCGAATTTGTTGGACACCTTACTACCTGTCGCCAGAGATCTGTCAGATTAACCC 540
QY 552 TACAACAAATAAAGGATATTGGTCTCTTGGCTGTGCTTATATGAGCTCTGCACTTT 611
D 541 TACAACAAATAAAGGATATTGGTCTCTTGGCTGTGCTTATATGAGCTCTGCACTTT 600
QY 612 AAACATCTTTTGGAGGTAAACATTCAGAGCTGGTTCTGAAAGATTGTGCAAGCAT 671
D 601 AAACATCTTTTGGAGGTAAACATTCAGAGCTGGTTCTGAAAGATTGTGCAAGCAT 660
QY 672 TTTGCCCAATATCTCCGGGGTTTTCTGTGAGCTCCATTTCTTATATCTCAGCTCTTT 731
D 661 TTTGCCCAATATCTCCGGGGTTTTCTGTGAGCTCCATTTCTTATATCTCAGCTCTTT 720
QY 732 CAAGTATCTCTCGAGAACGACATCCATAAATTCATTTTGAAGGCCCTTTTAGAG 791
D 721 CAAGTATCTCTCGAGAACGACATCCATAAATTCATTTTGAAGGCCCTTTTAGAG 780
QY 792 AATCTTATTTCCCAATATTTGACTCTCCTCAGGTCATTTCAGGAGAAATTCAGTCACAT 851
D 781 AATCTTATTTCCCAATATTTGACTCTCCTCAGGTCATTTCAGGAGAAATTCAGTCACAT 840
QY 852 ATATGAGAGAGAGAGCCGCCAGCTTCTCGACATGCTGGGAAGGTGGTCCAGAAAGTGA 911
```

Db 841 ATATGACAGCAGGAGCGGCGAGCTTCTCGACATGCTGGGAAGTGGTCCAGAAGTGTAAA 900
QY 912 ATACAAAAAGTGAGATTCGGGGAAGTCCCAACCAAGATCAAGGATATCTGTGCAATT 971
Db 901 ATACAAAAAGTGAGATTCAGGGAAGTCCCAACCAAGATCAAGGATATCTGTGCAATT 960
QY 972 AAAAGGAAGTCTATATGTCATAGAAATGAATGGAGCACACAGCTGGAGCCAGAGGCC 1031
Db 961 AAAAGGAAGTCTATATGTCATAGAAATGAATGGAGCACACAGCTGGAGCCAGAGGCC 1020
QY 1032 AGATCTATAAAATGATAGAAAGCCCAAAATTTGCTGTCTGTGTGACATTATGATTAT 1091
Db 1021 AGATCTATAAAATGATAGAAAGCCCAAAATTTGCTGTCTGTGTGACATTATGATTAT 1080
QY 1092 TATTATGCTCAACTTGATATGCTGAGGAGAGGCCCAACCAAGATGATACCTTAT 1151
Db 1081 TATTATGCTCAACTTGATATGCTGAGGAGAGGCCCAACCAAGATGATACCTTAT 1140
QY 1152 CCTCAAGAAATACTCGAGTTAGGATTACGGTTCAGAAACAGGAGCATGGTCCATCCCA 1211
Db 1141 CCTCAAGAAATACTCGAGTTAGGATTACGGTTCAGAAACAGGAGCATGGTCCATCCCA 1200
QY 1212 AGTCAATGGCTCTGAGTACCTTCAGAGAAATTTGAAGTCAACAAATATAGTTGAAA 1271
Db 1201 AGTCAATGGCTCTGAGTACCTTCAGAGAAATTTGAAGTCAACAAATATAGTTGAAA 1260
QY 1272 GTGAGAGCAATTTGGTCTTCTGCTCATCTTCTGCCGAGCCAAATTTACACACAGACAA 1331
Db 1261 GTGAGAGCAATTTGGTCTTCTGCTCATCTTCTGCCGAGCCAAATTTACACACAGACAA 1320
QY 1332 GAGCTAAGAGTAAATGGAGAAGCCTAGATTCCAGGAGCTGCATTAGGAAACAAACGAA 1391
Db 1321 GAGCTAAGAGTAAATGGAGAAGCCTAGATTCCAGGAGCTGCATTAGGAAACAAACGAA 1380
QY 1392 ATGAGGAACAGGAATATTTGGAAGCAGTTAGAGGAATATAGCAACAGTACCTCAATGAC 1451
Db 1381 ATGAGGAACAGGAGAACTCAAAATAAGTCAATAAAACCTATTGTGTGAAGAAGATAAC 1440
QY 1452 ATCAAGAAATTTAGAAAGAGATGGGAGAGAACCA-----GAGGACATTTGAAAAGAC 1505
Db 1441 CTGCCTGTCCATCAAGATGATCTGAGGAGAGACCTGTGCAGGACATTGAAAAGAC 1500
QY 1506 TTGAAACAAATGAGGCTTCAGAACACAAAGGAAGTAAATAATCCAGAACAGAAATATAA 1565
Db 1501 TTGAAACAAATGAGGCTTCAGAACACAAAGGAAGTAAATAATCCAGAACAGAAATATAA 1560
QY 1566 GCTAAGAGGGGTAAATTTGAAATTAATTTAGACAAATGATTTCTGATGAAAACATC 1625
Db 1561 GCTAAGAGGGGTAAATTTGAAATTAATTTAGACAAATGATTTCTGATGAAAACATC 1620
QY 1626 CTCGAAGAGGAGGCAATGGATATACCAATGAAACTTTGACCTTTGAGGATGGCATG 1685
Db 1621 CTCGAAGAGGAGGCAATGGATATACCAATGAAACTTTGACCTTTGAGGATGGCATG 1680
QY 1686 AAGTTTAAGGAATATGATGTGTAAGGAGCATGGAGATTATACAGACAAAGCATTTGAA 1745
Db 1681 AAGTTTAAGGAATATGATGTGTAAGGAGCATGGAGATTATACAGACAAAGCATTTGAA 1740
QY 1746 AAACCTTCACCTGCCGAGAGCAGCATTTACAGAACTGACT 1784
Db 1741 AAACCTTCACCTGCCGAGAGCAGGGTTTCCACGAGACT 1779

RESULT 6
US-09-940-921B-1
; Sequence 1, Application US/09940921B
; Patent No. US20020147320A1
; GENERAL INFORMATION:
; APPLICANT: Fiddle, Carl Johan
; APPLICANT: Hilbun, Erin
; APPLICANT: Nepomnichy, Boris
; APPLICANT: Hu, Yi

; TITLE OF INVENTION: No. US20020147320A1el Human Kinase Proteins and Polynucleotides E
; FILE REFERENCE: LEX-0227-USA
; CURRENT APPLICATION NUMBER: US/09/940,921B
; PRIOR FILING DATE: 2002-05-21
; PRIOR APPLICATION NUMBER: US 60/229,280
; PRIOR FILING DATE: 2000-08-31
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 2052
; TYPE: DNA
; ORGANISM: homo sapiens
US-09-940-921B-1

Query Match 69.8%; Score 1654; DB 9; Length 2052;
Best Local Similarity 94.5%; Pred. No. 0;
Matches 1763; Conservative 0; Mismatches 10; Indels 93; Gaps 1;
QY 12 ATGATAAGTACGATGTGATTAAAGCCATCGGCAAGTGCCTTCGGAAAGCATACTTA 71
Db 1 ATGATAAGTACGATGTGATTAAAGCCATCGGCAAGTGCCTTCGGAAAGCATACTTA 60
QY 72 GCTAAAGGGAATCAGATAGCAAGCAGTGTCTATAAAGAGATCAATTTGAAAAGATG 131
Db 61 GCTAAAGGGAATCAGATAGCAAGCAGTGTCTATAAAGAGATCAATTTGAAAAGATG 120
QY 132 CCCATACAAGAAAAAGAGCTTCAAGAAAAAGAGTGTCTTCTGGAAGAGATGAAACAT 191
Db 121 CCCATACAAGAAAAAGAGCTTCAAGAAAAAGAGTGTCTTCTGGAAGAGATGAAACAT 180
QY 192 CCCAACATGTAGGCTTCTCAATTCATTTCAAGAGAAATGGCAGGCTGTATTGTAATG 251
Db 181 CCCAACATGTAGGCTTCTCAATTCATTTCAAGAGAAATGGCAGGCTGTATTGTAATG 240
QY 252 GAATATTGTGATGAGGGGATCTCATGAAAAGATCAATAGACAAACGGGGTGTGTTATT 311
Db 241 GAATATTGTGATGAGGGGATCTCATGAAAAGAGTCAATAGACAAACGGGGTGTGTTATT 300
QY 312 AGTGAAGATCAGATCCTCGSTTGGTTTGTACAGATTTCTCTAGAGCTAATAACATATT 371
Db 301 AGTGAAGATCAGATCCTCGSTTGGTTTGTACAGATTTCTCTAGAGCTAATAACATATT 360
QY 372 GACAGGAAGATTATACACAGGACATAAAGCTCAGAACATTTTCTTAGCAGAACCGGA 431
Db 361 GACAGGAAGATTATACACAGGACATAAAGCTCAGAACATTTTCTTAGCAGAACCGGA 420
QY 432 ATGTTGGCAAGCTTGGGACTTTGGTATAGCAAGAGTCTCTGAAATAATTCATCGAACTT 491
Db 421 ATGTTGGCAAGCTTGGGACTTTGGTATAGCAAGAGTCTCTGAAATAATTCATCGAACTT 480
QY 492 GCTGGAATCTGTATTGGAACACCTTACTACCTGTCCCAGAGATCTGTCAGAAATAACCC 551
Db 481 GCTGGAATCTGTATTGGAACACCTTACTACCTGTCCCAGAGATCTGTCAGAAATAACCC 540
QY 552 TACAACATAAAACGGATATTGGTCTCTTGGCTGTGTCTTATATAGCTCTGCACACTT 611
Db 541 TACAACATAAAACGGATATTGGTCTCTTGGCTGTGTCTTATATAGCTCTGCACACTT 600
QY 612 AAACATCTTTTGGGGTAAACACTTACAGCAGCTGGTCTGGAAGATTTGTCAAGCAGAT 671
Db 601 AAACATCTTTTGGGGTAAACACTTACAGCAGCTGGTCTGGAAGATTTGTCAAGCAGAT 660
QY 672 TTTGCCCCAATATCTCGGGGGTTTTCTGTGAGGCTCCATTCCTTGATATCTCAGCTCTTT 731
Db 661 TTTGCCCCAATATCTCGGGGGTTTTCTGTGAGGCTCCATTCCTTGATATCTCAGCTCTTT 720
QY 732 CAAGTATCTCTCGAGACCGACCATCCATAAATTTCCATTTTGAAGGCCCTTTTAGAG 791
Db 721 CAAGTATCTCTCGAGACCGACCATCCATAAATTTCCATTTTGAAGGCCCTTTTAGAG 780
QY 792 AATCTTATTTCCAAATATTTGACTCCTCAGCTCAATCAGGAGAAATTCAGTCAGATGCTT 851
Db 781 AATCTTATTTCCAAATATTTGACTCCTCAGCTCAATCAGGAGAAATTCAGTCAGATGCTT 840

Matches 1215; Conservative 0; Mismatches 4; Indels 3; Gaps 1;			
Qy	1	CATTGGAGACCACTGGATAGTACGATGCTGATTAAGCCATCGGCGAAGCTGCTTCGGGA	60
Db	88	CATTGGAGACCACTGGATAGTACGATGCTGATTAAGCCATCGGCGAAGCTGCTTCGGGA	147
Qy	61	AAGCATACTTAGCTAAGGGAAATCAGATAGCAAGCACTGTGTCTATAAAGAGAGATCAATT	120
Db	148	AAGCATACTTAGCTAAGGGAAATCAGATAGCAAGCACTGTGTCTATAAAGAGAGATCAATT	207
Qy	121	TTGAAAAGATGCCATACAAAGAAAAGAGCTTCAAGAAAAGAGATGATCTTCTGGAAA	180
Db	208	TTGAAAAGATGCCATACAAAGAAAAGAGCTTCAAGAAAAGAGATGATCTTCTGGAAA	267
Qy	181	AGATGAAACATCCCAACATTTAGCTTCTTCAATTCATTTCAAGAGAAATGCGAGCTGT	240
Db	268	AGATGAAACATCCCAACATTTAGCTTCTTCAATTCATTTCAAGAGAAATGCGAGCTGT	327
Qy	241	TTATTGTAAATGGAATTTGTATGGAGGGATCTCATGAAAGGATCAATAGACACCGG	300
Db	328	TTATTGTAAATGGAATTTGTATGGAGGGATCTCATGAAAGGATCAATAGACACCGG	387
Qy	301	GTGTGTTATTAGTGAAGATCAGATCCTCGGTTGGTTGTACAGATTTCTTAGGACTAA	360
Db	388	GTGTGTTATTAGTGAAGATCAGATCCTCGGTTGGTTGTACAGATTTCTTAGGACTAA	447
Qy	361	AACATATTCATGACAGGAAGATATTACAGGGACATAAAGCTCAGAACATTTTCTTAA	420
Db	448	AACATATTCATGACAGGAAGATATTACAGGGACATAAAGCTCAGAACATTTTCTTAA	507
Qy	421	GCAAGACCGAATGGTGGCAAGCTTGGGACCTTGGTATACCAAGAGTCCTGAATATT	480
Db	508	GCAAGACCGAATGGTGGCAAGCTTGGGACCTTGGTATACCAAGAGTCCTGAATATT	567
Qy	481	CCATGGAACCTTGCTCGAATCTGTTATGGAACACCTTACTACCTGCCAGAGATCTGTC	540
Db	568	CCATGGAACCTTGCTCGAATCTGTTATGGAACACCTTACTACCTGCCAGAGATCTGTC	627
Qy	541	AGAATAAACCCCTACACAAATAAAGGATTTGGTCTCTTGGCTGTGTCTTATGAGC	600
Db	628	AGAATAAACCCCTACACAAATAAAGGATTTGGTCTCTTGGCTGTGTCTTATGAGC	687
Qy	601	TCGTGCACATTAAACATCCCTTTGAGGGTAACAACTTACAGACGCTGGTCTTGAAGATT	660
Db	688	TCGTGCACATTAAACATCCCTTTGAGGGTAACAACTTACAGACGCTGGTCTTGAAGATT	747
Qy	661	GTCAGACATTTTGGCCCAATATCTCCGGGGTTTTCTCGTGAAGTCCATTCCTTGATAT	720
Db	748	GTCAGACATTTTGGCCCAATATCTCCGGGGTTTTCTCGTGAAGTCCATTCCTTGATAT	807
Qy	721	CTCAGCTCTTCAAGTATCTCCTCGAGACCGACCATCCATTAATTCATTTTGAAGGC	780
Db	808	CTCAGCTCTTCAAGTATCTCCTCGAGACCGACCATCCATTAATTCATTTTGAAGGC	867
Qy	781	CTTTTGTAGAGATCTTATTTCCCAATATTTGACTCTCTGAGGTCA---TTCAGGAAGAT	837
Db	868	CTTTTGTAGAGATCTTATTTCCCAATATTTGACTCTCTGAGGTCA---TTCAGGAAGAT	927
Qy	838	TCAGTCACATGCTTATGACAGACGAGGCGCCAGCTTCTCGACATGCTGGGAAGTGG	897
Db	928	TCAGTCACATGCTTATGACAGACGAGGCGCCAGCTTCTCGACATGCTGGGAAGTGG	987
Qy	898	TCAGAGAGTGTAAATACAAAAGAGTGAGATTCCGGGGAAAGTGCCCAACCAAGATCAAGA	957
Db	988	TCAGAGAGTGTAAATACAAAAGAGTGAGATTCCGGGGAAAGTGCCCAACCAAGATCAAGA	1047
Qy	958	TATCTGTGCCAATTAAGGAATGCTTATATTGATGAAATGAATGAGACCAACAGCTG	1017
Db	1048	TATCTGTGCCAATTAAGGAATGCTTATATTGATGAAATGAATGAGACCAACAGCTG	1107
Qy	1018	GAGCCAGAGGCGCAGATCTATAAATGATAGAAAGCCCAAAATTCGTGCTGTG	1077
Db	1108	GAGCCAGAGGCGCAGATCTATAAATGATAGAAAGCCCAAAATTCGTGCTGTG	1167

RESULT 8	
US-10-114-270-185	
; Sequence 185 Application US/10114270	
; Publication No. US20040030110A1	
; GENERAL INFORMATION:	
; APPLICANT: Guo, Xiaojia	
; APPLICANT: Kekuda, Ramesh	
; APPLICANT: Miller, Charles E.	
; APPLICANT: Malyankar, Uriel M.	
; APPLICANT: SPYTEK, Kimberly A.	
; APPLICANT: Patturajan, Meera	
; APPLICANT: Liu, Zhaozhong	
; APPLICANT: Gusev, Vladimir Y.	
; APPLICANT: Li, Li	
; APPLICANT: Vernet, Corine	
; APPLICANT: zethusen, Bryan D.	
; APPLICANT: Gorman, Linda	
; APPLICANT: Shenoy, Suresh G.	
; APPLICANT: Pena, Carol E.A.	
; APPLICANT: Smithson, Glennda	
; APPLICANT: Burgess, Catherine E.	
; APPLICANT: Gerlach, Valerie	
; APPLICANT: Padigaru, Muralidhara	
; APPLICANT: Shinkets, Richard A.	
; APPLICANT: Gangolli, Esha A.	
; APPLICANT: Taupier Jr., Raymond J.	
; APPLICANT: Casman, Stacie J.	
; APPLICANT: Ji, Weizhen	
; APPLICANT: Anderson, David W.	
; APPLICANT: Liete, Mario W.	
; APPLICANT: Rastelli, Luca	
; APPLICANT: Edinger, Shlomit R.	
; APPLICANT: Stone, David J.	
; APPLICANT: MacDougall, John R.	
; APPLICANT: Rothenberg, Mark E.	
; TITLE OF INVENTION: No. US20040030110A1el Proteins and Nucleic Acids Encoding Same	
; FILE REFERENCE: 21402-322C	
; CURRENT APPLICATION NUMBER: US/10/114,270	
; CURRENT FILING DATE: 2002-11-27	
; PRIOR APPLICATION NUMBER: 60/281,086	
; PRIOR FILING DATE: 2001-04-03	
; PRIOR APPLICATION NUMBER: 60/281,136	
; PRIOR FILING DATE: 2001-04-03	
; PRIOR APPLICATION NUMBER: 60/281,863	
; PRIOR FILING DATE: 2001-04-05	
; PRIOR APPLICATION NUMBER: 60/281,906	
; PRIOR FILING DATE: 2001-04-05	
; PRIOR APPLICATION NUMBER: 60/282,020	
; PRIOR FILING DATE: 2001-04-06	
; PRIOR APPLICATION NUMBER: 60/282,930	
; PRIOR FILING DATE: 2001-04-10	
; PRIOR APPLICATION NUMBER: 60/282,934	
; PRIOR FILING DATE: 2001-04-10	
; PRIOR APPLICATION NUMBER: 60/283,512	
; PRIOR FILING DATE: 2001-04-12	
; PRIOR APPLICATION NUMBER: 60/283,710	
; PRIOR FILING DATE: 2001-04-13	
; PRIOR APPLICATION NUMBER: 60/284,234	
; PRIOR FILING DATE: 2001-04-17	

Qy	1078	GACATTATGATTATTATTATGCTCAACTTGATATGCTGAGGAGAGAGCCACAAACCAA	1137
Db	1168	GACATTATGATTATTATTATGCTCAACTTGATATGCTGAGGAGAGAGCCACAAACCAA	1227
Qy	1138	GTTATCACCCCTATTTCCTCAAGAAAATACTGGAGTTGAGGATTACGGTCAGGAAACGAGGC	1197
Db	1228	GTTATCACCCCTATTTCCTCAAGAAAATACTGGAGTTGAGGATTACGGTCAGGAAACGAGGC	1287
Qy	1198	ATGGTCCATCCCCCAAGTCAATG	1219
Db	1288	ATGGTCCATCCCCCAAGTCAATG	1309

; Remaining Prior Application data removed - See File Wrapper or PALM.

; NUMBER OF SEQ ID NOS: 470
 ; SEQ ID NO 185
 ; LENGTH: 1453
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 ; FEATURE:
 ; NAME/KEY: CDS
 ; LOCATION: (1)..(1198)
 US-10-114-270-185

Query Match 49.7%; Score 1176.8; DB 16; Length 1453;
 Best Local Similarity 99.1%; Pred. No. 5.3e-302;
 Matches 1197; Conservative 0; Mismatches 2; Indels 9; Gaps 1;

QY 12 ATGGATAAGTACGATGTGATTAAAGCCATCGGGCAAGGTGCTTCGGGAAAGCATCTTA 71
 DB 1 ATGGATAAGTACGATGTGATTAAAGCCATCGGGCAAGGTGCTTCGGGAAAGCATCTTA 60
 QY 72 GCTAAGGGAATCAGATAGCAAGCACTGTCTATTAAGAGATCAATTTTGAAGATG 131
 DB 61 GCTAAGGGAATCAGATAGCAAGCACTGTCTATTAAGAGATCAATTTTGAAGATG --- 117
 QY 132 CCCATACAGAAAAAGAGCTTCAAAGAAAGAGTGAATTTCTGGAAGATGAACAT 191
 DB 118 -----CAAGAAAAAGAGCTTCAAAGAAAGAGTGAATTTCTGGAAGATGAACAT 171
 QY 192 CCCACATGTAGCCCTTCTCAATTCATTTCAAGAGATGCGAGCTGTATTGTATG 251
 DB 172 CCCACATGTAGCCCTTCTCAATTCATTTCAAGAGATGCGAGCTGTATTGTATG 231
 QY 252 GAATATTGTGATGGAGGGATCTCATGAAAAAGGATCAATAGCAACGGGGTGTCTATT 311
 DB 232 GAATATTGTGATGGAGGGATCTCATGAAAAAGGATCAATAGCAACGGGGTGTCTATT 291
 QY 312 AGTGAAGTACAGTCCCTCGGTGGTGTGTAAGATTTCTTAGACCTAAACATATTTCAT 371
 DB 292 AGTGAAGTACAGTCCCTCGGTGGTGTGTAAGATTTCTTAGACCTAAACATATTTCAT 351
 QY 372 GACAGGAAGATATTACAGGGACATAAAGCTCAGAACATTTTCTTAGCAAGACGGA 431
 DB 352 GACAGGAAGATATTACAGGGACATAAAGCTCAGAACATTTTCTTAGCAAGACGGA 411
 QY 432 ATGTGGCAAGCTTGGGACTTTGGTATAGCAAGAGTCTCTGAATAATTCCATGGAACTT 491
 DB 412 ATGTGGCAAGCTTGGGACTTTGGTATAGCAAGAGTCTCTGAATAATTCCATGGAACTT 471
 QY 492 GCTCGAATTTGATGGACACCTTACTACCTGTCCCAGAGATCTGTCAAGATAAACCC 551
 DB 472 GCTCGAATTTGATGGACACCTTACTACCTGTCCCAGAGATCTGTCAAGATAAACCC 531
 QY 552 TACAACTATAAAGCGATATTGGTCTCTTGGCTGTCTTATATGAGCTCTGCACACTT 611
 DB 532 TACAACTATAAAGCGATATTGGTCTCTTGGCTGTCTTATATGAGCTCTGCACACTT 591
 QY 612 AAACATCTTTTGGGGTAAACACTTACAGAGCTGGTTCGAAGATTTGTCAAGCAAT 671
 DB 592 AAACATCTTTTGGGGTAAACACTTACAGAGCTGGTTCGAAGATTTGTCAAGCAAT 651
 QY 672 TTTGCCCCAATATCTCCGGGGTTTCTCGTGAAGTCCATTTCTTGATATCTCAGCTCTTT 731
 DB 652 TTTGCCCCAATATCTCCGGGGTTTCTCGTGAAGTCCATTTCTTGATATCTCAGCTCTTT 711
 QY 732 CAAGTATCTCTCGAGACCGACCAATCCATAAATTTCCATTTTGAAGAGCCCTTTTAGAG 791
 DB 712 CAAGTATCTCTCGAGACCGACCAATCCATAAATTTCCATTTTGAAGAGCCCTTTTAGAG 771
 QY 792 AATCTTATTTCCCAATATTTTCACTCTCGAGTCAATTCAGGAAGATTTTCAGTCACTGCTT 851
 DB 772 AATCTTATTTCCCAATATTTTCACTCTCGAGTCAATTCAGGAAGATTTTCAGTCACTGCTT 831
 QY 852 ATATGCAAGAGAGGAGCCAGCTTCTCGACATCTCTGGAGAGTGGTCCAGAAAGTGTAAA 911

DB 832 ATATGCAAGAGAGGAGCGCCAGCTTCTCGACATGTGTGGAAAGTGGTCCAGAAAGTGTAAA 891
 QY 912 ATACAAAAAGTGTGATTTCCGGGAAAGTCCCCACCAAGATCAAGGATATCTCTGCAATT 971
 DB 892 ATACAAAAAGTGTGATTTCCGGGAAAGTCCCCACCAAGATCAAGGATATCTCTGCAATT 951
 QY 972 AAAAGGAATGCTATATTGTCATAGAAATGAATGAGACCAACAGCTGGAGCCAGAGGCC 1031
 DB 952 AAAAGGAATGCTATATTGTCATAGAAATGAATGAGACCAACAGCTGGAGCCAGAGGCC 1011
 QY 1032 AGATCTATAAAATGATAGAAAGCCCAAAATTTGCTGTCTGTGTGGACATTATGATTAT 1091
 DB 1012 AGATCTATAAAATGATAGAAAGCCCAAAATTTGCTGTCTGTGTGGACATTATGATTAT 1071
 QY 1092 TATTATGCTCAACTTGTATATGCTGAGGAGAGAGCCCAACCAAGTATACCCCTATT 1151
 DB 1072 TATTATGCTCAACTTGTATATGCTGAGGAGAGAGCCCAACCAAGTATACCCCTATT 1131
 QY 1152 CCTCAAGAAATACCTGGAGTTTCTGAGGATACGGTCAAGAAACGAGCATGCTCATCCCA 1211
 DB 1132 CCTCAAGAAATACCTGGAGTTTCTGAGGATACGGTCAAGAAACGAGCATGCTCATCCCA 1191
 QY 1212 AGTCAATG 1219
 DB 1192 AGTCAATG 1199

RESULT 9

US-09-783-320-50
 ; Sequence 50, Application US/09783320
 ; Patent No. US20020038011A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Walke, D. Wade
 ; APPLICANT: Hu, Yi
 ; APPLICANT: Nepomnichy, Boris
 ; APPLICANT: Turner, C. Alexander Jr
 ; APPLICANT: Zambrowicz, Brian
 ; TITLE OF INVENTION: NO. US20020038011A1el Human Kinases and Polynucleotides Encoding
 ; FILE REFERENCE: LEX-0137-USA
 ; CURRENT APPLICATION NUMBER: US/09/783,320
 ; CURRENT FILING DATE: 2001-02-15
 ; PRIOR APPLICATION NUMBER: US 60/183,582
 ; PRIOR FILING DATE: 2000-02-18
 ; PRIOR APPLICATION NUMBER: US 60/184,014
 ; PRIOR FILING DATE: 2000-02-22
 ; NUMBER OF SEQ ID NOS: 50
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 50
 ; LENGTH: 5426
 ; TYPE: DNA
 ; ORGANISM: homo sapiens
 US-09-783-320-50

Query Match 15.6%; Score 369.4; DB 9; Length 5426;
 Best Local Similarity 66.0%; Pred. No. 8e-87;
 Matches 551; Conservative 0; Mismatches 281; Indels 3; Gaps 1;

QY 9 ACCATGGGATAGTACGATGTGATTAAAGGCCCATCGGGCAAGGTGCTTCGGGAAAGCATAC 68
 DB 531 ATCATGGAGAGTATGTTAGACTACAGAAAGATTGGAGAAGTTTCAATTTGGAAGCCATT 590
 QY 69 TTAGCTAAGGGAATCAGATAGCAGCACTGTGTCTATAAAGAGATCAATTTTGAAG 128
 DB 591 CTTGTAAATCTTACAGAAAGATGGCAGACAGTATGTATCAAGGAAATTAACATCTCAGA 650
 QY 129 ATGCCCATCAAGAAAAAGAGCTTCAAGAAAAAGAGTATTCTTCTGAAAAAGATGAAA 188
 DB 651 ATGTCCAGTAAGAAGAGAGAGATCAAGGAGAGAGTTGCAAGTATTTGGCAACATGAG 710
 QY 189 CATCCCAACATTGAGCCTTCTCAATTTCAAGAGATGCGAGCTGTTTATTGTA 248
 DB 711 CATCAAAATATTGTCCAGTATAGAGAAATCAATTTGAAGAAATGGCTCTCTCTACATAGTA 770

249 ATGGAATATTGTCATGGAGGGGATCTCATCAAAAGGATCAATAGACAACGGGTGTGTTA 308
771 ATGCAATTACTGTGAGGGAGGGGATCTGTTTAAGCGAATAAATGCTCAGAAAGCGGTTTG 830
309 TTTAGTGAAGATCAGATCTCGGTTGGTTGTACAGATTCTCTAGGACTAAACATATT 368
831 TTTCAAGAGATCAGATTGTTGAGCTGGTTGTACAGATATGTTGGCCCTGAAACATGTA 890
369 CATGACAGGAAGATATTACACAGGGGACATAAAAGCTCAGAACATTTTCTTAGCAAGAAC 428
891 CATGATAGAAAATTTCTCATCGAGACATTAATCTCAGACATATTTTAACTAAGAT 950
429 GGAATGGTGCAAAAGCTTGGGACTTGGTATAGCAAGAGTCTCTGATTAATTCATGGA 488
951 GGAACAGT---ACAACCTTGGAGATTTTGGAAATTTGCTAGAGTTCTTAATAGTACTGTAG 1007
489 CTTCCTCGAATTTGTAATGGAACACCTTACTCTGCCAGAGATCTGTGAGAAATAA 548
1008 CTGGCTCGAATTTGCAATAGGACCCCATCTACTTGTACCTGAAATCTGTGAAACAAA 1067
549 CCTACAAACATAAAACGGATATTGGTCTCTTGGCTGTGTCTTATATGAGCTCTGCACA 608
1068 CCTTACAATAATAAAAGTGACATTTGGGCTCTGGGGTGTGTCTTTATGAGCTGTGACA 1127
609 CTTAACAATCTTTGAGGGTAAACACTTACAGCAGCTGGTCTGAGAGATTTGTCAAGCA 668
1128 CTTAAAGAAATCCTTAGGATAGACCATCAGTCAACTCCATATCGAGAAAGGTTTATA 1307
789 GAGAACTTTATCCCAATATTTGACTCTCTGAGTCAATTCAGAGAAATTCAGTC 843
1308 GCCAAACGCAATGAAAGTTTCTCTCTCTCAGCTATTTCAGAGAAATTTGTGTC 1362

RESULT 10

US-10-037-270-246
; Sequence 246, Application US/10037270
; Publication No. US20030104529A1
; GENERAL INFORMATION:
; APPLICANT: Tang, Y. Tom
; APPLICANT: Liu, Chenghua
; APPLICANT: Asundi, Vinod
; APPLICANT: Zhang, Jie
; APPLICANT: Ren, Feiyan
; APPLICANT: Chen, Rui-hong
; APPLICANT: Zhao, Qing A.
; APPLICANT: Wehrman, Tom
; APPLICANT: Xue, Aigong J.
; APPLICANT: Yang, Yonghong
; APPLICANT: Wang, Jian-Rui
; APPLICANT: Zhou, Ping
; APPLICANT: Ma, Yunding
; APPLICANT: Wang, Dunrui
; APPLICANT: Wang, Zhiwei
; APPLICANT: Tillingshast, John
; APPLICANT: Drmanac, Radoje T.
; TITLE OF INVENTION: No. US20030104529A1el Nucleic Acids and
; FILE REFERENCE: Polypeptides
; CURRENT APPLICATION NUMBER: US/10/037,270
; CURRENT FILING DATE: 2002-01-04
; PRIOR APPLICATION NUMBER: 09/552,317
; PRIOR FILING DATE: 2000-04-25
; PRIOR APPLICATION NUMBER: 09/488,725
; PRIOR FILING DATE: 2000-01-21
; NUMBER OF SEQ ID NOS: 1104

; SOFTWARE: pt FL_genes Version 1.0
; SEQ ID NO 246
; LENGTH: 5448
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (534)..(4178)
US-10-037-270-246

Query Match 15.6%; Score 369.4; DB 15; Length 5448;
Best Local Similarity 66.0%; Pred. No. 8e-87;
Matches 551; Conservative 0; Mismatches 281; Indels 3; Gaps 1;

QY 9 ACCATGATAGTACGATGTGATTAAAGGCCATCGGCAAGGTGCTTTGGGAAAGCATAC 68
Db 531 ATCATGAGAGATGATTGTAGACATACAGAGATTGGAGAGGTTCAITTTGAAAAGCCATT 590
QY 69 TTACCTAAAGGAAATCAGATAGCAAGACTGTGTCTATAAAAGAGATCAATTTTGAAGAAG 128
Db 591 CTTGTTAAATCTCAGAGAGATGGCAGACAGTATGTTATCAAGGAAATTAACATCTCAAGA 650
QY 129 ATGCCCATACAGAAAAGAGCTTCAAGAAAAGAGTGTCTTCTGGAAGAAATCAAA 188
Db 651 ATGTCCAGTAAGAAAGAGAGATCAAGGAGAGAGTTGCAGTATTGGCAACATGAAG 710
QY 189 CATCCCAACATTTAGGCTTTCTTCAATTCATTTCAAGAGAAATGCGAGGCTGTTTATTGTA 248
Db 711 CATCCAAATATTGCCAGTATAGAGAAATCTTTGAAGAAATAGGCTCTCTCTACATAGTA 770
QY 249 ATGGAATATTGTGATGAGGGGATCTCATGAAAGAGTCAATAGACACGGGTGTGTTA 308
Db 771 ATGGAATATTGTGAGGAGGGGATCTGTTTAAAGCAATTAATGCTCAGAAAGGCGTTTG 830
QY 309 TTTAGTGAAGATCAGATCTCGGTTGGTTGTACAGATTTCTCTAGGACTAAACATATT 368
Db 831 TTTCAAGAGGATCAGATTTTGGACTGTTGTACAGATATGTTTGGCCCTGAAACATGTA 890
QY 369 CATCAGAGAGATATTACACAGGACATAAAAGCTCAGAACATTTTCTTAGCAAGAAC 428
Db 891 CATGATAGAAAATTTCTTCATCGAGACATTAATCTCAGAACATATTTTTAACTAAAGAT 950
QY 429 GGAATGGTGGCAAAAGCTTTGGTATAGCAAGAGTCTCTGAATAATTTCCATCGAA 488
Db 951 GGAACAGT---ACAACCTTGGAGATTTTGGAAATTTGCTAGAGTTCTTANTAGTACTGTAGAG 1007
QY 489 CTTCCTCGAATCTTATGGAACACCTTACTCTGTCCCGCAGAGATCTGTCAAGATAAA 548
Db 1008 CTGGCTCGAATCTGCATAGGGACCCCATACTACTTGTCACTGAAATCTGTGAAAACAAA 1067
QY 549 CCCTACACAAATAAAACGGATATTGGTCTCTTGGCTGTGTCTTATATGAGCTCTGCACA 608
Db 1068 CTTTACAATTAATAAAGTGAATTTGGGCTCTGGGGTGTGTCTTTTATGAGCTGTGTACA 1127
QY 609 CTTAAACATCTTTTGAAGGTAACACTTACAGCAGCTGGTTCTGGAAGATTTGTCAAGCA 668
Db 1128 CTTAAACATCTTTTGAAGCTGGCAGTATGAAAACCTGGTACTGGAAGATAATATCTGGA 1187
QY 669 CATTTTGGCCCCAATATCTCGGGGTTTCTCGTGAGGCTCCATCTCTTGATATCTCAGCTC 728
Db 1188 TCCTTTCCACCTGTGTCTTTGCAATTTCTTATGATCTCCGAGTTTGGGTCTCAGTTA 1247
QY 729 TTTCAAGTATCTCTCGAGACCGGACCATCCATAAATTTCCATTTTGAAGAGGCCCTTTTA 788
Db 1248 TTTAAAGAAATCCTAGGATAGACCATCAGTCACTCCATATTGGGAAAGGTTTATA 1307
QY 789 GAGAACTTTATCCCAATATTTGACTCTCTGAGTCAATTCAGAGAAATTCAGTC 843
Db 1308 GCCAAACGCAATGAAAGTTTCTCTCTCTCAGCTATTTCAGAGAAATTTGTGTC 1362

RESULT 11

US-10-117-722-246

QY	609	CTTAAACATCTCTTTTGAGGGTAAACA	CTTACAGCAGCTGGTTCTGAAGATTTTCTCAAGCA	688
DB	1128	CTTAAACATGCTCTTTTGAAGCTGGCAGATG	GAAGAAACCTGGTACTGAAGATAATATCTGGA	1187
QY	669	CATTTTGGCCCAATATCTCCGGGGTTTCT	CGTCAGCTCCATTCCTTGATATCTCAGCTC	728
DB	1188	TCCTTTCCACCTGTGTCTTTCATATTCCT	TATGATCTCCGAGTTTGGTGTCTCAGTTA	1247
QY	729	TTTCAAGTATCTCTCGAGACCGACCAT	TCCATAAATTTCCATTTTGAAGAGCCCTTTTA	788
DB	1248	TTTAAAGAAATCTCTAGGATAGACCAT	CAGTCACTCCATATTTGGAGAAAGGTTTATA	1307
QY	789	GAGATCTTATTCCTCCCAATATTTGACT	CTCCCTGAGTCAATCAGGAAGATTCAGTC	843
DB	1308	GCCAAACGATGAAAGTTTCTCTCTCCT	CAGCTTATTCGAGAAAGATTTTGTCTC	1362
RESULT 12				
US-10-037-270-245				
; Sequence 245, Application US/10037270				
; Publication No. US20030104529A1				
; GENERAL INFORMATION:				
; APPLICANT: Tang, Y. Tom				
; APPLICANT: Liu, Chenghua				
; APPLICANT: Asundi, Vinod				
; APPLICANT: Zhang, Jie				
; APPLICANT: Ren, Peiyan				
; APPLICANT: Chen, Rui-hong				
; APPLICANT: Zhao, Qing A.				
; APPLICANT: Wehrman, Tom				
; APPLICANT: Xue, Aidong J.				
; APPLICANT: Yang, Yonghong				
; APPLICANT: Wang, Jian-Rui				
; APPLICANT: Zhou, Ping				
; APPLICANT: Ma, Yungqing				
; APPLICANT: Wang, Dunrui				
; APPLICANT: Wang, Zhiwei				
; APPLICANT: Tillinghast, John				
; APPLICANT: Dmanac, Radoje T.				
; TITLE OF INVENTION: NO. US20030104529A1el Nucleic Acids and				
; TITLE OF INVENTION: Polypeptides				
; FILE REFERENCE: 784CIP2B				
; CURRENT APPLICATION NUMBER: US/10/037,270				
; CURRENT FILING DATE: 2002-01-04				
; PRIOR APPLICATION NUMBER: 09/552,317				
; PRIOR FILING DATE: 2000-04-25				
; PRIOR APPLICATION NUMBER: 09/488,725				
; PRIOR FILING DATE: 2000-01-21				
; NUMBER OF SEQ ID NOS: 1104				
; SOFTWARE: pt_FL_genes Version 1.0				
; SEQ ID NO 245				
; LENGTH: 5532				
; TYPE: DNA				
; ORGANISM: Homo sapiens				
; FEATURE:				
; NAME/KEY: CDS				
; LOCATION: (534)..(4262)				
US-10-037-270-245				
Query Match 15.6%; Score 369.4; DB 15; Length 5532;				
Best Local Similarity 66.0%; Pred. No. 8.e-87;				
Matches 551; Conservative 0; Mismatches 281; Indels 3; Gaps 1;				
QY	9	ACATGCTAAGTACGATGTGATTAAGCCATCGGCAAGTGCCTTCGGGAAGCATAC	68	
DB	531	ATCATGGAGAAGTATGTTAGACTACAGAAAGATTCGAGAAGTTCATTTGGAAAGCCATT	590	
QY	69	TTAGCTTAAAGGGAATTCAGATAGCAAGCACTGTGTCTATTAAGAGAGATCAATTTTGA	128	
DB	591	CTTGTTAAATCTACAGAGATGGCAGACATGTTATCAGAGAAATTAATCATCTCAGA	650	
QY	129	ATGCCCATCAAGAAAGAAAGCTTCAAGAAAGAAAGTGTCTTCTGGAAAGATGAA	188	
DB	651	ATGTCCAGTAAAGAAAGAAAGAAATCAAGGAGAGAAAGTTGCAGTATTGGCAAAATCAAG	710	
QY	189	CATCCCAATTTGACCTTCTTCATTTCAATTTCAAGAGATGCGAGCTGTTTATTGTA	248	
DB	711	CATCCAAATTTGTCAGTATAGAAATCAATTTGAAGAAATGGCTCTCTCATATAGTA	770	
QY	249	ATGGAATTTGTGATGGGGGATCTCATGAAAGGATCAATAGACAACCGGGTGTGTTA	308	
DB	771	ATGGATTACTGTGAGGGGGATCTGTTTAAAGCAATTAATGCTCAGAAAGGGCTTTTG	830	
QY	309	TTTAGTGAATCAGATCCTCGTTGGTTGTTGATAGATTTCTTAGGACTAAACATATT	368	
DB	831	TTTCAAGAGGATCAGATTTTGGACTGGTTTGTACAGATATGTTTGGCCCTGAAACATGTA	890	
QY	369	CATGACAGAGAATATTACAGAGGACATATAAAGCTCAGAACATTTTCTTAGCAAGAC	428	
DB	891	CATGATAGAAATTTCTTCATCGACATTAATCTCAGACATATTTTAACTAAAGAT	950	
QY	429	GGAATGGTGGCAAGCTTGGGACCTTGGTATAGCAAGAGTCTCGAATAATTTCCATCGAA	488	
DB	951	GGAAACAGT---ACAACCTTGGAGATTGGGAATTTGCTAGAGTCTTAAATAGTACTGTAGAG	1007	
QY	489	CTTGCTCGAACCTTGTATTGGAACCTTACTACCTGTCCCGCAGAGATCTGTCAGATATAA	548	
DB	1008	CTGGCTCGAACCTTGTATAGAGGACCCATCTACTTGTCACTTCACTTGTGAACACAA	1067	
QY	549	CCCTACAAATATAACCGATATTTGGTCTCTCTGGCTGTCTTATATGAGCTCTGCACA	608	
DB	1068	CCTTCAATATAAAGTACATTTGGGCTCTGGGGTGTCTCTTTATGAGCTGTGATCA	1127	


```
Db 651 ATGTCCAGTAAAGAAAGAGAAGAAATCAAGGAGAGAAGTTGCAGTATTGGCAACATCAAG 710
Qy 189 CATCCCAACATTTGAGCTCTTCAATTTCAATTTCAAGAGAAATGCGAGGCTGTTTATTGTA 248
Db 711 CATCCAAATATTGTCAGATATAGAGATCAATTTGAAGAAATGCGCTCTCTACATAGTA 770
Qy 249 ATGGAATATTGTGATGAGGGGATCTCATGAAAGAGATCAATAGACAACGGGTGTGTTA 308
Db 771 ATGGAATTACTGTGAGGGAGGGATCTGTTTAAGCGAATAAATGCTCAGAAAGCGTTTGG 830
Qy 309 TTTAGTGAAGATCAGATCTCGGTGTTGTTGTTACAGATTTCTAGGATTAACACATATT 368
Db 831 TTTCAAGAGATCAGATTTGGACTGTTGTTGTTACAGATATTGTTGCCCTGAAACATGA 890
Qy 369 CATGACAGGAAGATATTACACAGGGACATAAAAGCTCAGAACATTTTTTCTTACAGAAAC 428
Db 891 CATGATAGAAAATTTCTTCATCGAGACATTAATCTCAGAACATATTTTAACTAAAGAT 950
Qy 429 GGAATGGTGCAAGCTTGGGACTTTGGTATAGCAAGAGTCTGTAATTAATCCATGGAA 488
Db 951 GGAACAGT---ACAACCTGGAGATTTTGGAAATGCTAGAGTTCTTAAATAGTACTGTAG 1007
Qy 489 CTGTCTCGAACTTGTATTGGAACACCTTACTACTCTGCCAGAGATCTGTCAAGATAAA 548
Db 1008 CTGGCTCGAACTTGCATAGGACCCCATACTACTTGTCACTGAAATCTGTGAAACAAA 1067
Qy 549 CCTACAAACATAAACGGATATTGCTCTTTGGTGTGTTGTTATATAGCTCTGCACA 608
Db 1068 CTTTACAAATAAATAAGTGACATTTGGGCTCTGGGCTGTGCTCTTTATGAGCTGTGTACA 1127
Qy 609 CTTTAAACATCCTTTGAGGGTAAACATTAACAGCAGCTGTTCTGAGGATTTGTCACAGA 668
Db 1128 CTTTAAACATGCTTTGAGCTGCGAGTATGAAACCTGTACTGAAGATAATATCTGGA 1187
Qy 669 CATTTTGCCCAATATCTCGGGGTTTCTCGTGAAGCTCCATTCCTTGATATCTCAGCTC 728
Db 1188 TCCTTTCCACCTGTCTTTGCAATTTCTCTATGATCTCGGCAAGTTTGGTGTCTCAGTTA 1247
Qy 729 TTTCAAGTATCTCTCGAGACCCCATCCATTAATTTCCATTTTGAAGGCCCTTTTATA 788
Db 1248 TTTAAAGAAATCCTAGGATAGACCATCAGTCAACTCCATATTTGAGAAAGGTTTATA 1307
Qy 789 GAGAATCTTATTTCCCAATATTGACTCTCTGAGCTCAATTCAGAGAAATTCAGTC 843
Db 1308 GCCAAAGCATTTGAAAGTTTCTCTCTCTCAGCTTATTGCAGAGAAATTTGTC 1362
```

RESULT 13

```
US-10-117-722-245
; Sequence 245, Application US/10117722
; Publication No. US20030219744A1
; GENERAL INFORMATION:
; APPLICANT: Tang, Y. Tom
; APPLICANT: Liu, Chenghua
; APPLICANT: Asundi, Vinod
; APPLICANT: Zhang, Jie
; APPLICANT: Dmanac, Radoje T.
; TITLE OF INVENTION: Polypeptides
; FILE REFERENCE: 784CIP25CIP
; CURRENT APPLICATION NUMBER: US/10/117,722
; CURRENT FILING DATE: 2002-04-04
; PRIOR APPLICATION NUMBER: 09/620,312
; PRIOR FILING DATE: 2000-07-19
; PRIOR APPLICATION NUMBER: 09/552,317
; PRIOR FILING DATE: 2000-04-25
; PRIOR APPLICATION NUMBER: 09/488,725
; PRIOR FILING DATE: 2000-01-21
; NUMBER OF SEQ ID NOS: 1104
; SOFTWARE: pt_FL_genes Version 1.0
; SEQ ID NO 245
; LENGTH: 5532
```

```
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (534)..(4262)
US-10-117-722-245
```

Query Match 15.6%; Score 369.4; DB 15; Length 5532;

Best Local Similarity 66.0%; Pred. No. 8.1e-87;

Matches 551; Conservative 0; Mismatches 281; Indels 3; Gaps 1;

```
Qy 9 ACCATGGATAGTACGATGTGATTAAGCCATCGGCAAGGTGCTTCGGGAAAGCATAC 68
Db 531 ATCATGGAGAAGTATGTTAGACTACAGAAGATTGGAGAAGGTTCAATTGGAAAAAGCCATT 590
Qy 69 TTAGCTAAAGGGAATCAGATAGCAAGCACCTGTCTATAAAGAGATCAATTTTGAAGA 128
Db 591 CTGTGTTAATCTACAGAAGATGGCAGACAGTATGTTATCANGGAATTAATCTCAGA 650
Qy 129 ATGCCCATACAGAAAAAGAGCTTTCAAGAAAGAGTATCTTCTGGAAAAAGATGAAA 188
Db 651 ATGTCAGTAAAGAAAGAGAGAATAATCAAGGAGAGAAGTTGCAGTATTGGCAAAATGAAG 710
Qy 189 CATCCCAACATTTAGCCCTTCTTCAATTTCAAGAGATGCGAGCTCTGTTATTGTA 248
Db 711 CATCAAAATATTGCCAGTATAGAGATCAATTTGAAGAAATGGCTCTCTCTCATAGTA 770
Qy 249 ATGGAATATTGTGAGGGGATCTCATGAAAAAGGATCAATAGACAACGGGGTGTGTTA 308
Db 771 ATGGATTTACTGTGAGGGAGGGGATCTGTTTAAAGCAATAAATGCTCAGAAAGGCGTTTGG 830
Qy 309 TTTAGTGAAGATCAGATCTCGGTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 368
Db 831 TTTCAAGAGGATCAGATTTTGGACTGTTTGTACAGATATGTTTGGGCTCTTAAATAGTACTGTAG 890
Qy 369 CATGACAGGAAGATATTACACAGGGACATAAAAGCTCAGAACATTTTCTTAGCAAGAAC 428
Db 891 CATGATAGAAAATTTCTTCATCGAGACATTAATCTCAGAACATATTTTAACTAAAGAT 950
Qy 429 GGAATGGTGGCAAGCTTGGGACTTTGGTATAGCAAGAGTCTGTAATTAATCCATGGAA 488
Db 951 GGAACAGT---ACAACCTGGAGATTTTGGAAATGCTAGAGTTCTTAAATAGTACTGTAG 1007
Qy 489 CTGTCTCGAACTTGTATTGGAACACCTTACTACTCTGCCAGAGATCTGTCAAGATAAA 548
Db 1008 CTGGCTCGAACTTGCATAGGACCCCATACTACTTGTCACTGAAATCTGTGAAACAAA 1067
Qy 549 CCTACAAACATAAACGGATATTGCTCTTTGGTGTGTTGTTATATAGCTCTGCACA 608
Db 1068 CTTTACAAATAAATAAGTGACATTTGGGCTCTGGGCTGTGCTCTTTATGAGCTGTGTACA 1127
Qy 609 CTTTAAACATCCTTTGAGGGTAAACATTAACAGCAGCTGTTCTGAGGATTTGTCACAGA 668
Db 1128 CTTTAAACATGCTTTTGAAGCTGGCAGTATGAAACCTGTACTGAAGATAATATCTGGA 1187
Qy 669 CATTTTGCCCAATATCTCGGGGTTTCTCGTGAAGCTCCATTCCTTGATATCTCAGCTC 728
Db 1188 TCCTTTCCACCTGTCTTTGCAATTTCTCTATGATCTCGGCAAGTTTGGTGTCTCAGTTA 1247
Qy 729 TTTCAAGTATCTCTCGAGACCCCATCCATTAATTTCCATTTTGAAGGCCCTTTTATA 788
Db 1248 TTTAAAGAAATCCTAGGATAGACCATCAGTCAACTCCATATTTGAGAAAGGTTTATA 1307
Qy 789 GAGAATCTTATTTCCCAATATTGACTCTCTGAGCTCAATTCAGGAAAGATTCAGTC 843
Db 1308 GCCAAAGCATTTGAAAGTTTCTCTCTCTCAGCTTATTGCAGAGAAATTTGTC 1362
```

RESULT 14

```
US-10-618-941-29
; Sequence 29, Application US/10618941
; Publication No. US20040197792A1
; GENERAL INFORMATION:
```

; APPLICANT: WHYTE, DAVID
; APPLICANT: MANNING, GERARD
; APPLICANT: CAENEPEEL, SEAN
; TITLE OF INVENTION: NOVEL KINASES
; FILE REFERENCE: 034536-0321
; CURRENT APPLICATION NUMBER: US/10/618,941
; CURRENT FILING DATE: 2003-07-15
; PRIOR APPLICATION NUMBER: 60/395,632
; PRIOR FILING DATE: 2002-07-15
; NUMBER OF SEQ ID NOS: 143
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 29
; LENGTH: 5583
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-618-941-29

Query Match 15.6%; Score 369.4; DB 18; Length 5583;
Best Local Similarity 66.0%; Pred. No. 8.1e-87; Mismatches 281; Indels 3; Gaps 1;
Matches 551; Conservative 0; Mismatches 281; Indels 3; Gaps 1;
QY 9 ACCATGGATAAGTACGATGTGATTAAGGCCATCGGCAAGGTGCTTCGGGAAAGCATAC 68
Db |||||
QY 490 ATCATGGAGAGTAGTGTAGACTACAGAAGATTGGAGAAGTTCATTTGGAAGGCCATT 549
Db |||||
QY 69 TTAGCTAAAGGGAATCAGATAGCAGCACTGTGTCTATAAAGAGATCAATTTGAAAG 128
Db |||||
QY 550 CTTGTAAATCTACAGAAGATGGCAGACAGTATGTTATCAAGGAAATTAACATCTCAAG 609
Db |||||
QY 129 ATGCCCATACAGAAAAGAGCTTCAAGAAAGAGTGAATCTCTGGAAGATCAAAA 188
Db |||||
QY 610 ATGTCAGTAAAGAAAGAGAGATCAAGGAGAGAGTTCAGATTTGGCNAACATGAAG 669
Db |||||
QY 189 CATCCCAACATTTAGTACCTCTTCAATTCATTTCAAGAGATGCGAGCTGTTTATTGTA 248
Db |||||
QY 670 CATCCAAATATGTCACAGTATAGAGAAATCAATTTGAAGAAATGCTCTCTACATAGTA 729
Db |||||
QY 249 ATGGATATTGATGGAGGGGATCTCATGAAGAGATCAATACAGCAACGGGGTGTTA 308
Db |||||
QY 730 ATGGATTACTGTGGAGGGGATCTGTTTAAGCGAATAAATGCTCAGAAGGGGTTTG 789
Db |||||
QY 309 TTTAGTGAAGATCAGATCTCGTTTGGTTGTACAGATTTCTTAGGACTAAACATATT 368
Db |||||
QY 790 TTTCAAGAGGATCAGATTTGSACTGTTGTACAGATATGTTTGGCCCTGAAACATGTA 849
Db |||||
QY 369 CATGACAGAGATATTACAGGACATAAAGCTCAGACATTTTCTTAGCAGTAAGAC 428
Db |||||
QY 850 CATGATAGAAAATTCCTCATCGAGACATTAATCTCAGAACATATTTTAACTAAGAT 909
Db |||||
QY 429 GGAATGGTGGCAAGCTTGGGACCTTGGTATAGCAAGAGTCTCGAATAATTCATCGAA 488
Db |||||
QY 910 GGAACAGT--ACAACCTGGAGATTGGAAATGCTAGAGTCTTAATAGTACTGTAGAG 966
Db |||||
QY 489 CTTGCTGAACTGTATTGGAACACTTACTACTCTGTCGCCAGAGATCTGTCAGAAATAA 548
Db |||||
QY 967 CTTGGCTGAACTGTGATAGGACCCCATACTTGTCACTGTAATCTGTGAAACAAA 1026
Db |||||
QY 549 CCTACACAATAAAGCGATATTTGGTCTCTGGCTGTGCTTATATAGCTCTGCACA 608
Db |||||
QY 1027 CTTTAAATATATAAAGTGACATTTGGGCTCTGGGGTGTGCTTTTATAGCTGTGTACA 1086
Db |||||
QY 609 CTTAAACATCTTTTGGGGTAAACACTTACAGCAGCTGGTTCTGAAGATTGTCAAGCA 668
Db |||||
QY 1087 CTTAAACATGCTTTTGAAGCTGGCAGTATGAAAACCTGTGACTGAAGATAATATCTGGA 1146
Db |||||
QY 669 CATTTTGCCCAATATCTCCGGGTTTCTCGTGAGCTCCATTCCTTGATATCTCAGCTC 728
Db |||||
QY 1147 TCTTTTCCACCTGTGCTCTTGTGATTTATTCCTATGATCTCCGAGTTTGGTGTCTCAGTTA 1206
Db |||||
QY 729 TTTCAAGTATCTCTCGAGACCCGACCATCCATAAATTCATTTTGAAGGCCCTTTTAA 788
Db |||||
QY 1207 TTTAAAGAAATCCTTAGGATAGACCATCAGTCAACTCCATATTGGAGAAAGGTTTATA 1266
Db |||||

QY 789 GAGAACTTTATTCCCAATATTGACTCTCTGAGGTCAATTCAGGAGAAATTCAGTC 843
Db |||||
QY 1267 GCCAAACGCATTGAAAAGTTTCTCTCTCTCAGCTATTTCGAGAGAAATTTGTC 1321
Db |||||

RESULT 15
US-09-783-320-3
; Sequence 3, Application US/09783320
; Patent No. US20020038011A1
; GENERAL INFORMATION:
; APPLICANT: Walke, D. Wade
; APPLICANT: Hu, Yi
; APPLICANT: Nepomnichy, Boris
; APPLICANT: Turner, C. Alexander Jr
; APPLICANT: Zambrowicz, Brian
; TITLE OF INVENTION: NO. US20020038011A1 Human Kinases and Polynucleotides Encoding
; FILE REFERENCE: LEX-0137-USA
; CURRENT APPLICATION NUMBER: US/09/783,320
; CURRENT FILING DATE: 2001-02-15
; PRIOR APPLICATION NUMBER: US 60/183,582
; PRIOR FILING DATE: 2000-02-18
; PRIOR APPLICATION NUMBER: US 60/184,014
; PRIOR FILING DATE: 2000-02-22
; NUMBER OF SEQ ID NOS: 50
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 3645
; TYPE: DNA
; ORGANISM: homo sapiens
US-09-783-320-3

Query Match 15.5%; Score 368; DB 9; Length 3645;
Best Local Similarity 66.0%; Pred. No. 1.5e-86;
Matches 549; Conservative 0; Mismatches 280; Indels 3; Gaps 1;

QY 12 ATGGATAAGTACGATGTGATTAAGGCCATCGGCAAGTGGCTTCGGGAAAGCATACTTA 71
Db |||||
QY 1 ATGGAGAAGTATGTTAGACTACAGAAGATTGGAGAAGTTCATTTGGAAGGCCATCTCT 60
Db |||||
QY 72 GCTAAAGGAATCAGATAGCAAGCACTGTGTCATAAAGAGATCAATTTGAAAGATG 131
Db |||||
QY 61 GTTAAATCTACAGAAGATGGCAGACAGTATGTTATCAAGGAAATTAACATCTCAAGATG 120
Db |||||
QY 132 CCATACAGAAAAGAAAGCTTCAAGAAAGAAAGTGAATCTTCTGGAAGAAAGTGAACAT 191
Db |||||
QY 121 TCCAGTAAAGAAAGAGAGAATCAAGGAGAGAAGTTGCAGTATTGGCAACATGAAGCAT 180
Db |||||
QY 192 CCCAATTTGTAGCCCTTCTCAATTCNTTCAAGAGATGGCAGGCTGTTTATTGTAATG 251
Db |||||
QY 181 CCAATATTGTCAGTATAGAGAATCATTTGAAGAAATGGCTCTCTCTACATAGTAATG 240
Db |||||
QY 252 GAATATTGTAGTGGAGGGATCTCATGAAAGAGATCAATAGACAACGGGGTGTGTTATT 311
Db |||||
QY 241 GATTACTGTGAGGAGGGATCTGTTTAAGCGAAATAATGCTCAGAAGGGCTTTGTTT 300
Db |||||
QY 312 AGTGAAGATCAGATCCTGGTTGGTTTGTACAGATTTCTTAGAGCTAAACATATTGAT 371
Db |||||
QY 301 CAAGAGGATCAGATTTGGACTGGTTTGTACAGATATGTTGGCCCTGAAACATGTACAT 360
Db |||||
QY 372 GACAGAAGATATTACAGAGGACATAAAGCTCAGAACATTTTCTTAGCAAGAACGA 431
Db |||||
QY 361 GATAGAAAATTTCTTCATCGACACATTAATCTCAGAACATATTTTAACTAAGATGGA 420
Db |||||
QY 432 ATGGTGGCAAGCTTGGGACTTTGGTATAGCAAGAGTCTCGAATAATTCATGGAATCT 491
Db |||||
QY 421 ACAGT--ACAACCTGGAGATTGGAATTTGTAGAGTCTTAATAGTACTGTAGAGCTG 477
Db |||||
QY 492 GCTCGAAGTGTATTGGAACACCTTACTACCTGTCCAGAGATCTGTCAGATAAACC 551
Db |||||
QY 478 GCTCGAAGTGTGATAGGACCCCATACTACTGTCACTGAAATCTGTGAAACAACT 537
Db |||||
QY 552 TACAACAATAAAGCGATATTTGGTCTCTTGGCTGTGCTTATATGAGCTCTGCACTT 611
Db |||||

Db	538	TACAATAATAAAAGTGACATTTGGGCTCTGGGTGTGTCTTTATGAGCTGTGTACACTT	597
Qy	612	AAACATCCTTTTGGAGGTAACTTACAGCAGCTGTTCTGAAGATTTGTCAAGCACAT	671
Db	598	AAACATGCTTTTGAAGCTGSCAGTATGAATAAACCTGTTACTGAAGATAATATCTGGATCT	657
Qy	672	TTTGCCCCAATATCTCGGGGTTTCTCGTGAGCTCCATTCCTTGATATCTCAGCTCTTT	731
Db	658	TTTCCACCTGTGTCTTTGCAATTATCTATGATCTCGGAGTTTGGTGTCTCAGTTATTT	717
Qy	732	CAAGTATCTCTCGAGACCGACCATCCATAAATTCATTTTGAAGGCCCTTTTAGAG	791
Db	718	AAAGAAATCCTAGGATAGACCATCAGTCACTCCATATTGGAGAAAGTTTATAGCC	777
Qy	792	AATCTTATTCCTCCAAATATTTGACTCCTGAGGTCAATTCAGGAAGAAATTCAGTC	843
Db	778	AAAGCATTGAAAAGTTTCTCTCTCTCAGCTTATTGCAGAGAAATTTGTC	829

Search completed: November 8, 2004, 05:40:12
 Job time : 4632 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 7, 2004, 21:12:58 ; Search time 6658 Seconds
(without alignments)
12971.187 Million cell updates/sec

Title: US-10-730-010-1
Perfect score: 2370
Sequence: 1 cattgagaccatgataag.....ctgtcatgcgcttaagagcg 2370

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 32822875 seqs, 18219865908 residues
Total number of hits satisfying chosen parameters: 65645750

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : EST:*
1: gb_est1:*
2: gb_est2:*
3: gb_hta:*
4: gb_est3:*
5: gb_est4:*
6: gb_est5:*
7: gb_est6:*
8: gb_gsa1:*
9: gb_gsa2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1218.4	51.4	1483	CR598570	full-length
2	842.2	35.5	890	CR598570	full-length
3	825.6	34.8	980	CR598570	full-length
4	597.4	25.2	1434	CR598570	full-length
5	597.4	25.2	2407	CR598570	full-length
6	594	25.1	603	CR598570	full-length
7	574.8	24.3	872	CR598570	full-length
8	547.8	23.1	670	CR598570	full-length
9	516.8	21.8	928	CR598570	full-length
10	485	20.5	856	CR598570	full-length
11	464.8	19.6	603	CR598570	full-length
12	455.4	19.2	603	CR598570	full-length
13	452.4	19.1	659	CR598570	full-length
14	446.2	18.8	678	CR598570	full-length
15	392	16.5	639	CR598570	full-length
16	391.6	16.5	643	CR598570	full-length
17	369.4	15.6	1628	CR598570	full-length
18	364.6	15.4	857	CR598570	full-length
19	355.8	15.0	2070	CR598570	full-length
20	355.8	15.0	5362	CR598570	full-length
21	354.2	14.9	3568	CR598570	full-length
22	347	14.6	3103	CR598570	full-length
23	329.8	13.9	648	CR598570	full-length
24	320.2	13.5	904	CR598570	full-length

25	313.4	13.2	732	7	CN531454
26	309.6	13.1	706	6	CD802687
27	308	13.0	630	4	BJ029013
28	293.8	12.4	718	6	CB312741
29	274.2	11.6	700	7	CF531621
30	265.2	11.2	2465	3	AK045202
31	259.4	10.9	763	7	CN160628
32	258.4	10.9	498	5	EX342706
33	254.4	10.7	649	6	CD291255
34	248.8	10.5	860	5	BP142906
35	248.2	10.5	812	5	BU322582
36	244.2	10.3	913	5	EQ432111
37	242.4	10.2	908	5	EX375816
38	241.2	10.2	927	5	BU516066
39	241	10.2	907	6	CA790699
40	240.6	10.2	730	7	CN531024
41	239	10.1	1971	3	CR612915
42	238.8	10.1	819	6	CF285034
43	238	10.0	930	5	EX364188
44	238	10.0	1835	3	CR606903
45	235.4	9.9	769	6	CF148064

ALIGNMENTS

RESULT 1 CR598570 1483 bp mRNA linear HTC 21-JUL-2004
LOCUS full-length cDNA clone CSODL005YK18 of B cells (Ramos cell line)
DEFINITION Cot 25-normalized of Homo sapiens (human).

ACCESSION CR598570 GI:50479377
VERSION HTG; CNSLUT_CDNA.

KEYWORDS Homo sapiens (human)
SOURCE Homo sapiens

ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 1483)
AUTHORS Li,W.B., Gruber,C., Jesse,J. and Polayes,D.

TITLE Full-length cDNA libraries and normalization
JOURNAL Unpublished
REMARK Contact : Feng Liang Email : fliang@lifetech.com URL : http://fulllength.invitrogen.com/ invitrogen Corporation 1600 Paraday Avenue

REFERENCE 2 (bases 1 to 1483)
AUTHORS Genoscope.
TITLE Direct Submission
JOURNAL Submitted (20-JUL-2004) Genoscope - Centre National de Sequencage : BP 191 91006 EVRY cedex - FRANCE (E-mail : segref@genoscope.cns.fr

COMMENT - Web : www.genoscope.cns.fr
1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-strand cDNA was digested with Not I and cloned into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library was normalized. Library was constructed by Life Technologies, a division of Invitrogen.
FEATURES Location/Qualifiers
source 1..1483
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CSODL005YK18"
/tissue_type="B cells (Ramos cell line) Cot 25-normalized"
/plasmid="pCMVSPORT_6"

ORIGIN
Query Match 51.4%; Score 1218.4; DB 3; Length 1483;
Best Local Similarity 99.9%; Pred. No. 52-294;
Matches 1219; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Oy 1 CATTGGAGACCATGTAAGTACGATGATTAAAGCCATCGGCGAGGTGCCTTCGGGA 60
Bj061491 BJ061491
Db 83 CATTGGAGACCATGTAAGTACGATGATTAAAGCCATCGGCGAGGTGCCTTCGGGA 142

```

QY 61 AAGCATCTTAGCTAAAGGGAATCAGATAGCAGCACTGTGTCATATAAAGAGATCAATT 120
DB 143 AAGCATCTTAGCTAAAGGGAATCAGATAGCAGCACTGTGTCATATAAAGAGATCAATT 202
QY 121 TTGAAAAGATCCCATACAGAAAAGAGAGCTTCAAAAGAAAAGAGTGAATCTTCTCGAAA 180
DB 203 TTGAAAAGATCCCATACAGAAAAGAGAGCTTCAAAAGAAAAGAGTGAATCTTCTCGAAA 262
QY 181 AGATGAACATCCCAACGATCTGTAGCTTCTTCAATTCATTCAGAGATGGCAGCTGT 240
DB 263 AGATGAACATCCCAACGATCTGTAGCTTCTTCAATTCATTCAGAGATGGCAGCTGT 322
QY 241 TTATTTGAATGAATTTGTGATGGAGGGATCTCATGAAAGAGATCAATAGACACGGG 300
DB 323 TTATTTGAATGAATTTGTGATGGAGGGATCTCATGAAAGAGATCAATAGACACGGG 382
QY 301 GTGTGTTATTTAGTGAAGATCAGATCTCTCGGTGGTTGTTGTACAGATTTCTTAGACTAA 360
DB 383 GTGTGTTATTTAGTGAAGATCAGATCTCTCGGTGGTTGTTGTACAGATTTCTTAGACTAA 442
QY 361 AACATATTCATGACAGAGATATTACAGGGGACATAAAGGCTCAGAACATTTTCTTA 420
DB 443 AACATATTCATGACAGAGATATTACAGGGGACATAAAGGCTCAGAACATTTTCTTA 502
QY 421 GCAGAACGGGAATGTTGGCAAGCTTGGGACATTTGGTATAGCAAGAGTCTCGAATAAT 480
DB 503 GCAGAACGGGAATGTTGGCAAGCTTGGGACATTTGGTATAGCAAGAGTCTCGAATAAT 562
QY 481 CCATGGAACTTCTCGACTTGTATTGGACACACCTTACTACCTGTCCCAAGAGATCTGC 540
DB 563 CCATGGAACTTCTCGACTTGTATTGGACACACCTTACTACCTGTCCCAAGAGATCTGC 622
QY 541 AGATAAACCCTACAAATAAAGCGATATTGGTCTCTTGGGTGTCTTATATGAGC 600
DB 623 AGATAAACCCTACAAATAAAGCGATATTGGTCTCTTGGGTGTCTTATATGAGC 682
QY 601 TCTGCACATTTAAACATCTTTGAGGGTAAACATTTACAGCAGCTGGTCTGAGATTT 660
DB 683 TCTGCACATTTAAACATCTTTGAGGGTAAACATTTACAGCAGCTGGTCTGAGATTT 742
QY 661 GTCAAGACATTTTGCCTCCATATCTCCGGGTTTCTCGTGAGCTCCATCTCTTGATAT 720
DB 743 GTCAAGACATTTTGCCTCCATATCTCCGGGTTTCTCGTGAGCTCCATCTCTTGATAT 802
QY 721 CTCAGCTCTTTCAAGATATCTCTCGAGACCGACCATCCATAAATTCATTTGAAAGGC 780
DB 803 CTCAGCTCTTTCAAGATATCTCTCGAGACCGACCATCCATAAATTCATTTGAAAGGC 862
QY 781 CCTTTTGAAGAATCTTATCCCAATATTTGACTCTCGAGGTCATTCAGGAAGATTTCA 840
DB 863 CCTTTTGAAGAATCTTATCCCAATATTTGACTCTCGAGGTCATTCAGGAAGATTTCA 922
QY 841 GTCACATCTTATATGACAGACGAGCGCAGCTTCTCGACATGTCTGGAAAGTGTCTC 900
DB 923 GTCACATCTTATATGACAGACGAGCGCAGCTTCTCGACATGTCTGGAAAGTGTCTC 982
QY 901 AGAAGTGTAATAACAAAAGTGAGATTCGGGGAAGTCCCAACCAAGATCAGGATAT 960
DB 983 AGAAGTGTAATAACAAAAGTGAGATTCGGGGAAGTCCCAACCAAGATCAGGATAT 1042
QY 961 CTGTGCCAATTTAAAGGAATCTTATTTGCAATAGAAATGAATGGAGACCCAGCTGGAG 1020
DB 1043 CTGTGCCAATTTAAAGGAATCTTATTTGCAATAGAAATGAATGGAGACCCAGCTGGAG 1102
QY 1021 CCCAGAGCCAGATCTTATAAATATGATAGAAAGACCAAAATTTGCTGTCTGTGGAC 1080
DB 1103 CCCAGAGCCAGATCTTATAAATATGATAGAAAGACCAAAATTTGCTGTCTGTGGAC 1162
QY 1081 ATTTATGATTTATTTATCTCAACTTGTATGTCTGAGGAGAGAGCCCAAAACCAAGTT 1140
DB 1163 ATTTATGATTTATTTATCTCAACTTGTATGTCTGAGGAGAGAGCCCAAAACCAAGTT 1222

1141 ATCACCTTATCTCTCAGAAAATATCTGGAGTTGAGGATTACGGTCAGGAAACGAGGCATG 1200
1223 ATCACCTTATCTCTCAGAAAATATCTGGAGTTGAGGATTACGGTCAGGAAACGAGGCATG 1282
1201 GTCCATCCCAAGTCAATGG 1220
1283 GTCCATCCCAAGTCAATGG 1302

RESULT 2
BX350584/c
LOCUS BX350584 Homo sapiens B CELLS (RAMOS CELL LINE) COT 25-NORMALIZED
DEFINITION Homo sapiens cDNA clone CS0DL005YK18 3-PRIME, mRNA sequence.
ACCESSION BX350584
VERSION BX350584.1 GI:30379530
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
MAMMALIA; EUTHERIA; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 890)
Li, W.B., Gruber, C., Jessee, J. and Polayes, D.
Full-length cDNA libraries and normalization
Unpublished (2001)
CONTACT: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 EVRY cedex - France
Email: segr@genoscope.cns.fr, Web : www.genoscope.cns.fr
1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime
end enriched double-strand cDNA was digested with Not I and cloned
into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library
was normalized. Library was constructed by Life Technologies, a
division of Invitrogen. This sequence belongs to sequence cluster
6942.r,
and it belongs to a clone representative of this cluster. For more
information about this cluster and the virtual cDNA, see
http://www.genoscope.cns.fr/cdna?s=CS0BAI043ZAI1_CS04113_1&c=6942.r

FEATURES
source
1..890
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0DL005YK18"
/cell_type="B CELLS (RAMOS CELL LINE) COT 25-NORMALIZED"
/clone_lib="Homo sapiens B CELLS (RAMOS CELL LINE) COT
25-NORMALIZED"
/note="1st strand cDNA was primed with a NotI-oligo(dT)
primer. Five prime end enriched, double-strand cDNA was
digested with Not I and EcoR V sites of the pCMVSPORT 6
vector. Library was normalized."

ORIGIN
Query Match 35.5%; Score 842.2; DB 5; Length 890;
Best Local Similarity 98.2%; Pred. No. 9.5e-200;
Matches 872; Conservative 0; Mismatches 14; Indels 2; Gaps 2;

QY 265 GAGGGATCTCATGAAAAGGATCAATAGACACGGGTGTGTTATTAGTCAAGATCAGA 324
DB 890 GAGGGATCTCATGAAAAGGATCAATAGACACGGGTGTG-TATTAAGTGAAGATCAGA 832
QY 325 TCTCGGTGGTTTGTACAGATTTCTTAGGACTAAACATATTCATGACAGGAGATAT 384
DB 831 TCTCGGTGGTTTGTACAGATTTCTTAGGACTAAACATATTCATGACAGGAGATAT 772
QY 385 TACACAGGACATAAAGCTCAGAACATTTTCTTAGCAAGACGGAATGTGCAAGC 444
DB 771 TACACAGGACATAAAGCTCAGAACATTTTCTTAGCAAGACGGAATGTGCAAGC 712
QY 445 TTGCGGACTTTGGTATAGCAAGAGTCTCGAATAATTCATGGAACCTGCTCGAACTGTA 504
```

```
Db 711 TTGGGACCTTGGTAGCAAGAGCTCTGAATAATTCATGGAACCTTGTCTCGAATTGTA 652
QY 505 TTGGAAACACCTTACTACCTGTCCCGAGAGATCTGCAGAAATAAACCCCTACAAATAAAA 564
Db 651 TTGGAAACACCTTACTACCTGTCCCGAGAGATCTGCAGAAATAAACCCCTACAAATAAAA 592
QY 565 CGGATATTTTGGTCTCTTGGCTGTGCTTATATAGCTCTGCACACTTAAACATCTTTTG 624
Db 591 CGGATATTTTGGTCTCTTGGCTGTGCTTATATAGCTCTGCACACTTAAACATCTTTTG 532
QY 625 AGGTTAAACACTTACAGAGCTGGTGTGAAGATTGTCAAGCAATTTTGCCTCCAT 684
Db 531 AGGTTAAACACTTACAGAGCTGGTGTGAAGATTGTCAAGCAATTTTGCCTCCAT 472
QY 685 CTCGGGGTTTCTCGTAGCTCCATCTCTGATATCTCAGCTCTTCAAGTATCTCTC 744
Db 471 CTCGGGGTTTCTCGTAGCTCCATCTCTGATATCTCAGCTCTTCAAGTATCTCTC 412
QY 745 GAGACCGACCATCCATAATTCATTTGAAAAGGCCCTTTTACAGAAATCTTATTC 804
Db 411 GAGACCGACCATCCATAATTCATTTGAAAAGGCCCTTTTACAGAAATCTTATTC 352
QY 805 ATATTTGACTCTGAGTCAATCAGGAAGATTGAGTCAATGCTTATATGAGAGAG 864
Db 351 ATATTTGACTCTGAGTCAATCAGGAAGATTGAGTCAATGCTTATATGAGAGAG 292
QY 865 GAGCCCGAGCTCTCGACATGCTGGGAAGTGTGTCAGAGTGTAAATACAAAAGTGA 924
Db 291 GAGCCCGAGCTCTCGACATGCTGGGAAGTGTGTCAGAGTGTAAATACAAAAGTGA 232
QY 925 GATTCGGGGAAAGTGTCCCAACAGATCAAGGATATCTGTGCCAATTTAAAGGAATGTA 984
Db 231 GATTCGGGGAAAGTGTCCCAACAGATCAAGGATATCTGTGCCAATTTAAAGGAATGTA 172
QY 985 TATTCATAGAAATGAATGGAGACACCAAGCTGGAGCCCGAGAGCCAGATCTATAAAA 1044
Db 171 TATTCATAGAAATGAATGGAGACACCAAGCTGGAGCCCGAGAGCTATAAAA 112
QY 1045 TGATAGAAAGACCCAAATTTGCTGTCTGTGGACATTTATGATTTATTTATGCTCAAC 1104
Db 111 TGATAGAAAGACCCAAATTTGCTGTCTGTGGACATTTATGATTTATTTATGCTCAAC 52
QY 1105 TTGATATCTGAGGAGGAGAGCCCAACCAAGTATACCCCTATTC 1152
Db 51 TTGATATGCT-AGGAGGAGAGCCCAACCAAGTATACCCCTATTC 5
```

RESULT 3

```
EX342707
LOCUS EX342707 Homo sapiens B CELLS (RAMOS CELL LINE) COT 25-NORMALIZED
DEFINITION Homo sapiens cDNA clone CS0DL005YK18 5-PRIME, mRNA sequence.
ACCESSION EX342707
VERSION EX342707.2 GI:46279550
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 980)
Li.W.B., Gruber,C., Jessee,J. and Polayes,D.
Full-length cDNA libraries and normalization
Unpublished (2001)
On May 2, 2003 this sequence version replaced gi:30344086.
Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 EVRY cedex - France
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime
end enriched, double-strand cDNA was digested with NotI and cloned
into the NotI and EcoRV sites of the pCMVSPORT 6 vector. Library
was normalized. Library was constructed by Life Technologies, a
division of Invitrogen. This sequence belongs to sequence cluster
```

6942.r,
and it belongs to a clone representative of this cluster. For more
information about this cluster and the virtual cDNA, see
<http://www.genoscope.cns.fr/cdna?s=CS0DL005BF09P1&c=6942.r>.

FEATURES

source

1..980
Location/Qualifiers
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0DL005YK18"
/cell_type="B CELLS (RAMOS CELL LINE) COT 25-NORMALIZED"
/cell_line="RAMOS CELL LINE"
/clone_lib="Homo sapiens B CELLS (RAMOS CELL LINE) COT
25-NORMALIZED"
/notes="1st strand cDNA was primed with a NotI-oligo(dT)
primer. Five prime end enriched, double-strand cDNA was
digested with NotI and cloned into the NotI and EcoRV
sites of the pCMVSPORT 6 vector. Library was normalized."

ORIGIN

Query Match 34.8%; Score 825.6; DB 5; Length 980;
Best Local Similarity 97.8%; Pred. No. 1.4e-195;
Matches 874; Conservative 5; Mismatches 9; Indels 6; Gaps 4;
QY 1 CAITGGAGACCATGGATTAAGTACGATCTGATTAAAG-GCCATCGGCAAGGTGCTTCGGG 59
Db 83 CAITGGAGACCATGGATTAAGTACGATCTGATTAAAGGCCATCGGCAAGGTGCTTCGGG 142
QY 60 AAAGCATACTTACGTAAGGGAAATCAGATAGCAAGACTGTGTCTATAAAGAGATCAAT 119
Db 143 AAAGCATACTTACGTAAGGGAAATCAGATAGCAAGACTGTGTCTATAAAGAGATCAAT 202
QY 120 TTTGAAAAGATGCCCATACAAAGAAAAGAGCTTCAAGAAAAGAGAGTATTCTTCGGAA 179
Db 203 TTTGAAAAGATGCCCATACAAAGAAAAGAGCTTCAAGAAAAGAGAGTATTCTTCGGAA 262
QY 180 AAGATGAAACATCCCAACATTTAGGCTTCTTCAATTCATTTCAAGAGAAATGCGAGCTG 239
Db 263 AAGATGAAACATCCCAACATTTAGGCTTCTTCAATTCATTTCAAGAGAAATGCGAGCTG 322
QY 240 TTTATTGTATGAAATATTGTGATGGAGGGATCTCATGAAAGGATCAATAGACAACGG 239
Db 323 TTTATTGTATGAAATATTGTGATGGAGGGATCTCATGAAAGGATCAATAGACAACGG 382
QY 300 GGTGTGTTATTATTAGTGAAGATCAGATCCTCGGTGTGTTGTACAGATTTCTCTAGACTA 359
Db 383 GGTGTGTTATTATTAGTGAAGATCAGATCCTCGGTGTGTTGTACAGATTTCTCTAGACTA 442
QY 360 AAACATATTTCATGACAGGAAGATATTACACAGGACATATAAGCTCAGAACATTTTCTT 419
Db 443 AAACATATTTCATGACAGGAAGATATTACACAGGACATATAAGCTCAGAACATTTTCTT 502
QY 420 AGCAAGAACGGAATGGTGGCAAGCTTGGGACTTTGGTATAGCAAGAGTCTCTGAATAAT 479
Db 503 AGCAAGAACGGAATGGTGGCAAGCTTGGGACTTTGGTATAGCAAGAGTCTCTGAATAAT 562
QY 480 TCCATGGAACCTTCTCGAACTTGTATTGGAACACCTTACTACCTGTCCCGAGATCTGT 539
Db 563 TCCATGGAACCTTCTCGAACTTGTATTGGAACACCTTACTACCTGTCCCGAGATCTGT 622
QY 540 CAGAATAAACCCCTACAAACAATAAAACGGATATTGTGCTCTTGGCTGTCTTATATGAG 599
Db 623 CAGAATAAACCCCTACAAACAATAAAACGGATATTGTGCTCTTGGCTGTCTTATATGAG 682
QY 600 CTCGACACATTAACATCTTTTGGAGGTAAACAATTACAGAGCTGGTCTTGAAGATT 659
Db 683 CTCGACACATTAACATCTTTTGGAGGT-ACAACCTTACAGAGCTGGTCTTGAAGATT 741
QY 660 TGTACAGACATTTTCCCGCAATATCTCCGGGGTTTCTCGTAGAGCTCCTTCTTGATA 719
Db 742 TGTCAAGACATTTTCCCGCAATATCTCCGGGGTTTCTCGTAGAGCTCCTTCTTGATA 801
QY 720 TCTCAGCTCTTTCAAGATATCTCTCCGAGACCGACCATCCATAAATTCCTATTTTGAAGG 779

Db 802 TCTCAGCTCTTTCAGTATCTCTCGAGACGACCAATCCATAAATTCATTTTGAAGAAG 861
 Qy 780 CCCTTTTATAGAGATCTTATTCACAAATATTTGACTCTCTGAGTCTCATTCAGGAAGATTC 839
 Db 862 CCCTTTTATAGAGATCTTATTCACAAATATTTGACTCTCTGAGTCTCATTCAGGAAG-ATTC 920
 Qy 840 AGTCACATGCTTATATCAGACGAGGAGCGCCAGCTTCTCAGATCTGCGGAAG 893
 Db 921 AGTCACATGCTTATATCAG---AGAGGAGGCGCAATTTCTCAGATCTGCGGAGG 971

RESULT 4
 AK054168 1434 bp mRNA linear HTC 03-APR-2004
 LOCUS AK054168.1 GI:26344040
 DEFINITION Mus musculus 2 days pregnant adult female oviduct cDNA, RIKEN
 full-length enriched library, clone:E230024112 product:weakly
 similar to protein kinase nek1 (EC 2.7.1.-) [Mus musculus], full
 insert sequence.
 ACCESSION AK054168
 VERSION AK054168.1
 KEYWORDS HTC; CAP trapper.
 SOURCE Mus musculus (house mouse)
 ORGANISM Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 1 Carninci, P. and Hayashizaki, Y.
 -Carninci, P. and Hayashizaki, Y.
 High-efficiency full-length cDNA cloning
 Meth. Enzymol. 303, 19-44 (1999)
 JOURNAL 99279253
 MEDLINE
 PUBMED 10349636
 REFERENCE 2
 AUTHORS Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K.,
 Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.
 TITLE Normalization and subtraction of cap-trapper-selected cDNAs to
 prepare full-length cDNA libraries for rapid discovery of new genes
 JOURNAL Genome Res. 10 (10), 1617-1630 (2000)
 MEDLINE 20499374
 PUBMED 11042159
 REFERENCE 3
 AUTHORS Shibata, K., Itoh, M., Aizawa, K., Nagao, K., Sasaki, N., Carninci, P.,
 Konno, H., Akiyama, J., Nishi, K., Katsunai, T., Tashiro, H., Itoh, M.,
 Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishine, T., Harada, A.,
 Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K.,
 Fujiwara, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watahiki, M.,
 Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsura, S., Kawai, J.,
 Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y.
 TITLE RIKEN integrated sequence analysis (RISA) system--384-format
 sequencing pipeline with 384 multichannel capillary sequencer
 JOURNAL Genome Res. 10 (11), 1757-1771 (2000)
 MEDLINE 20530913
 PUBMED 11076861
 REFERENCE 4
 AUTHORS The RIKEN Genome Exploration Research Group Phase II Team and the
 FANTOM Consortium.
 TITLE Functional annotation of a full-length mouse cDNA collection
 JOURNAL Nature 409, 685-690 (2001)
 REFERENCE 5
 AUTHORS The FANTOM Consortium and the RIKEN Genome Exploration Research
 Group Phase I & II Team.
 TITLE Analysis of the mouse transcriptome based on functional annotation
 of 60,770 full-length cDNAs
 JOURNAL Nature 409, 563-573 (2002)
 REFERENCE 6
 AUTHORS Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Bono, H., Carninci, P.,
 Fukuda, S., Furuno, M., Hanagaki, T., Hara, A., Hashizume, W.,
 Hayashida, K., Hayashizaki, Y., Hiramoto, K., Hiraoka, T., Hirozane, T.,
 Hori, F., Imorani, K., Ishii, Y., Itoh, M., Kagawa, I., Kasukawa, T.,
 Koyohji, H., Kawai, J., Kojima, Y., Kondo, S., Konno, H., Kouda, M.,
 Koya, S., Kurihara, K., Matsuyama, T., Miyazaki, A., Murata, M.,
 Nakamura, M., Nishikawa, K., Nomura, K., Numazaki, R., Ohno, M., Ohsato, N.,
 Okazaki, Y., Saito, R., Saitoh, H., Sakai, C., Sakai, K., Sakazume, N.,

Sano, H., Sasaki, D., Shibata, K., Shinagawa, A., Shiraki, T.,
 Sogabe, Y., Tagami, M., Tagawa, A., Takahashi, F., Takaku-Akahira, S.,
 Takeda, Y., Tanaka, T., Tomaru, A., Toya, T., Yasunishi, A.,
 Muramatsu, M. and Hayashizaki, Y.
 TITLE Direct Submission
 JOURNAL Submitted (16-JUL-2001) Yoshihide Hayashizaki, The Institute of
 Physical and Chemical Research (RIKEN), Laboratory for Genome
 Exploration Research Group, RIKEN Genomic Sciences Center (GSC),
 RIKEN Yokohama Institute, 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama,
 Kanagawa 230-0045, Japan (E-mail: genome-res@gsr.riken.jp,
 URL: http://genome.gsc.riken.jp/, Tel: 81-45-503-9222,
 Fax: 81-45-503-9216)
 COMMENT cDNA library was prepared and sequenced in Mouse Genome
 Encyclopedia Project of Genome Exploration Research Group in Riken
 Genomic Sciences Center and Genome Science Laboratory in RIKEN.
 Division of Experimental Animal Research in Riken contributed to
 prepare mouse tissues.
 Please visit our web site for further details.
 URL: http://genome.gsc.riken.jp/
 URL: http://fantom.gsc.riken.jp/.
 FEATURES
 Location/Qualifiers
 1..1434
 /organism="Mus musculus"
 /mol_type="mRNA"
 /strain="C57BL/6J"
 /db_xref="FANTOM.DB:E230024112"
 /db_xref="taxon:10090"
 /clone="E230024112"
 /sex="female"
 /tissue_type="oviduct"
 /clone_lib="RIKEN full-length enriched mouse cDNA library"
 /dev_stage="2 days pregnant adult"
 104..1114
 /note="unnamed protein product; putative
 weakly similar to protein kinase nek1 (EC 2.7.1.-) [Mus
 musculus] (PIR|S25284, evidence: PASTY, 46.5%ID,
 54.%length, match=1247)"
 /codon_start=1
 /protein_id="BAC35677.1"
 /db_xref="GI:26344041"
 /translation="MDNFHFKIIEGTFGKVLAKDSBSSHCIVKEISLTKESAS
 KNEVLLARMEHPNIVTFSSFOENGLFIVMEYCDGDLQRIORGVWFSBDQIL
 CWFQVLSGLKHIDRLKIHEDIKSONIFLSKNGWAKLGDGFTARTINDSMELAOIC
 AGTYILSPICQNRPNKNDIWSLGCVLVELCTLKHPPESNHFHLVXICQGRVA
 PISPHFRDQSLIPQFVRVSPQDRPSVTSLLKPFLETLIARLPIEVCSSRLQSSHA
 HMEENNAITGPTACWVSPWMDHVTWLSKMRGLGVVAIRLPLSLHKLPEFPFQCHKNE
 Kf"
 ORIGIN
 Query Match 25.2%; Score 597.4; DB 3; Length 1434;
 Best Local Similarity 80.3%; Pred. No. 2.6e-138; Indels 13; Gaps 2;
 Matches 730; Conservative 0; Mismatches 166; CATTGGGAAGGCTACT 161
 102 CCATGGATAAGTACGATGTGATTAAAGCCATCGGCGAAGTGCCTTCGGGAAGCACTACT 69
 70 TAGCTAAAGGGAATCAGATAGCAAGCACTGTGTCATAAAGAGATCAATTTGAAAGA 129
 162 TGGCTAAAGATAAATCAGAAAGCACTCACTGTGTCTATAAAGAAATCAGTTTGACAAG- 220
 130 TGCCCATACATAGAAAAGAAAGCTTCAAAAGAAAGATGATTCTTCTGGAAGATGAAC 189
 221 -----GAAAAGAGGCTCAAGACAGAGTATTCTTCTGGTAGGATGGAGC 269
 190 ATCCCAACATGTAGCTTCTTCAATTCATTTCAAGAGATGGCAGGCTGTTTATTGTAA 249
 270 ATCCCAATATCGTAACTTCTTTCAGCTGTTTCAAGAGAACGCGAGGCTGTTTATTGTAA 329
 250 TGGATATTTGATGGAGGATCTCATGAAAAGATCAATAGACAACGGGGCTGTGTAT 309
 330 TGGAAATCTGTGAGGAGGATCTCATGACAGAGATCCAGAGGCGGAGGATGATGT 389

```
QY 310 TTAGTGAATCAGATCCTCGGTGGTTGTGACAGATTTCTCAGAGTAAACATATTTC 369
Db 390 TCAGCGAAGACCAAGATCCTGTGTGGTTGTACAGATTTCTCAGAGTAAACATATTTC 449
QY 370 ATGACAGAGATATTACACAGGACGATAAAGCTCAGACATTTTCTTAGCAAGACG 429
Db 450 ATGACAGAGATATTACACAGGACGATAAAGCTCAGACATTTTCTTAGCAAGATG 509
QY 430 GAATGTGGCAAGCTTGGGACCTTGGTATAGCAAGCTCCTGAATAATTCATGGAAC 489
Db 510 GAATGTGGTGGCAAGCTTGGGACCTTGGGACGAAAGACACTGAATGACTCCATGGAAC 569
QY 490 TTGCTGAACTTGTATTGGAACCTTACTACCTGTCCAGAGATCTGCAAGATAAAC 549
Db 570 TTGCTGAACTTGTATTGGAACCTTACTACCTGTCCAGAGATCTGCAAGACAGC 629
QY 550 CCTACAAATAAAACGGATTTGGTCTCTTGGCTGTGTCTTATATAGAGCTCTGCACAC 609
Db 630 CATACAAATAAAACGGACATCTGGTCTCTTGGCTGTGTCTTATATAGAGCTCTGCACAC 689
QY 610 TTAACATCTTTGAGGTTAAACCTTACAGAGCTGTTCTGAAGATTTGTCAGACAC 669
Db 690 TCAGATCTCTTTGAGAGCAACACTTCCACATCTGTTCTGAAGATTTGTCAGAGAC 749
QY 670 ATTTGCGCCCAATATCTCGGGGTTTTCTCGTGAGCTCCATCTCTTGATATCTCAGCTCT 729
Db 750 GTGTTGCTCCATATCAACCCACTTCTCTCGTGACCTACAGTCTTGTATATAGCTCTGCACAC 809
QY 730 TTCAAGTATCTCTCAGACCGACCATCCATAAATTCATTTGAAAGGCCCTTTTAG 789
Db 810 TCAGATGTCTCTCAGACCGGCCCATCGGTTACGTCCTTTTGAAGACCCCTTTTAG 869
QY 790 AGAATCTTATCCCAATATTTGACTCTCGTGAGTCCATCTCTTGATATCTCAGCTCT 848
Db 870 AACTCTTATGCGCATCTTGTATCTCGTGAGTCTGTTCAAGAGAAATCCAGTCCCGACG 929
QY 849 CTTATATGAGACGACGCGCCAGCTTCTCGACATCTGTCGGAAGTGTCTCAGAGAGTGT 908
Db 930 CTCACATGAGAACATGCGCATTTGGCCACAGCTTTGTTGGAGAGTTAGTCCATGCTGG 989
QY 909 ABAATACAA 917
Db 990 ATCATGTAA 998

RESULT 5
AK032672
LOCUS
DEFINITION
Mus musculus 10 days neonate cerebellum cDNA, RIKEN full-length
enriched library, clone:653041J22 product:weakly similar to
protein kinase nek1 (EC 2.7.1.-) [Mus musculus], full insert
sequence.
ACCESSION
AK032672
VERSION
HTC; CAP trapper.
KEYWORDS
Mus musculus (house mouse)
SOURCE
Mus musculus
ORGANISM
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1
Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K.,
Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.
Normalization and subtraction of cap-trapper-selected cDNAs to
prepare full-length cDNA libraries for rapid discovery of new genes
Genome Res. 10 (10), 1617-1630 (2000)
2
20499374
PUBMED
11042159
```

REFERENCE
AUTHORS

3 Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P., Konno, H., Akiyama, J., Nishi, K., Kitsuami, T., Tashiro, H., Itoh, M., Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishine, I., Harada, A., Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K., Fujiwara, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watahiki, M., Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsuyama, S., Kawai, J., Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A., and Hayashizaki, Y. RIKEN integrated sequence analysis (RISA) system--384-format sequencing pipeline with 384 multipicapillary sequencer Genome Res. 10 (11), 1757-1771 (2000)

TITLE
JOURNAL
MEDLINE
PUBMED

20530913
11076861

REFERENCE
AUTHORS

4 The RIKEN Genome Exploration Research Group Phase II Team and the FANTOM Consortium.

TITLE
JOURNAL
MEDLINE
PUBMED

Nature 409, 585-590 (2001)

REFERENCE
AUTHORS

5 The FANTOM Consortium and the RIKEN Genome Exploration Research Group Phase I & II Team.

TITLE
JOURNAL
MEDLINE
PUBMED

Nature 420, 563-573 (2002)

REFERENCE
AUTHORS

6 (bases 1 to 2407)
Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Bono, H., Carninci, P., Fukuda, S., Furuno, M., Hanagaki, T., Hara, A., Hashizume, W., Hayashida, K., Hayatsu, N., Hiramoto, K., Hiraoka, T., Hirozane, T., Hori, F., Imotani, K., Ishii, Y., Itoh, M., Kagawa, I., Kasukawa, T., Katoh, H., Kawai, J., Kojima, Y., Kondo, S., Konno, H., Kouda, M., Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Murata, M., Nakamura, M., Nishi, K., Nomura, K., Numazaki, R., Ohno, M., Ohsato, N., Okazaki, Y., Saito, R., Saitoh, H., Sakai, C., Sakai, K., Sakazume, N., Sano, H., Sasaki, D., Shibata, K., Shinagawa, A., Shiraki, T., Sogabe, Y., Tagami, M., Tagawa, A., Takahashi, F., Takaku-Akahira, S., Takeda, Y., Tanaka, T., Tomaru, A., Toya, T., Yasunishi, A., Muramatsu, M. and Hayashizaki, Y.

TITLE
JOURNAL
MEDLINE
PUBMED

Submitted (16-JUL-2001) Yoshihide Hayashizaki, The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute; 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan (E-mail: genome-res@gsc.riken.jp, URL: http://genome.gsc.riken.jp/, Tel: 81-45-503-9222, Fax: 81-45-503-9216)

COMMENT

cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues.

FEATURES
source

URL: http://genome.gsc.riken.jp/
URL: http://fantom.gsc.riken.jp/
Location/Qualifiers
1. 2407

FEATURES
source

/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6J"
/db_xref="FANTOM,DB:653041J22"
/db_xref="taxon:10090"
/clone="653041J22"
/tissue_type="cerebellum"
/clone_lib="RIKEN full-length enriched mouse cDNA library"
/dev_stage="10 days neonate"
92. _1936

FEATURES
source

/note="unnamed protein product; putative weakly similar to protein kinase nek1 (EC 2.7.1.-) [Mus musculus] (PIR|S25284, evidence: FASTA, 46.5% ID, 54.7% length, match=1247)"
/codon_start=1
/protein_id="BAC27980.1"
/db_xref="GI:26328483"
/translation="MDNFLIKIIGGTGKVLAKDKSESSHCVKEISLTKEAKS

KNEVILLARMEHNI1VTFSSFOENGR1FVMEYCDGDLNOR1QORGVMEFSEQ1IL
 CWFVQISLGHK1HDK1LHRD1KSONI1FLSKNGMVA1K1GDFETARTLND1SMELAQTC
 AGPY1LSPIR1CONR1PNK1TD1WS1GCV1LYEL1CT1KHP1ESN1HEHL1VL1K1CQGRVA
 P1SPH1FSR1Q1OS1LI1PQ1LFR1VSP1QDR1SV1TS1LQ1R1F1ET1LI1AR1SL1Y1P1E1VC1SR1Q1SHA
 H1MEN1A1G1PT1AC1WR1SP1SA1YL1OR1K1FE1AQ1Y1KL1V1ER1OL1G1R1P1SV1PH1P1NE1G1KL1Q
 S1HWE1T1F1Q1B1L1Q1Y1K1N1K1MD1Q1E1Y1K1Q1E1I1R1Q1Y1H1ND1K1E1I1K1K1M1G1R1EL1K1V1Y1F1E1S
 LDK1C1S1E1D1V1Q1N1E1A1W1K1LN1AT1LS1P1E1D1G1T1F1O1E1H1R1C1K1E1H1E1Y1D1T1D1E1A1F1E1L1C1G1P1E1A1E
 G1F1Q1D1V1T1A1E1R1R1Q1W1D1A1G1A1P1H1T1L1R1M1A1D1V1T1S1C1T1P1W1D1D1G1O1V1T1M1E1G1S1V1E1N1G1K1Q
 W1L1D1V1P1T1P1C1A1A1E1C1A1C1S1G1S1A1S1K1G1E1T1W1K1P1K1E1Q1K1E1I1A1T1G1M1V1D1E1Q1L1E1P
 G1S1D1E1D1D1Q1E1R1G1N1K1N1S1P1S1A1T1K1I1A1S1P1C1M1A1C1S1N1S1V1P1P1L1Q1K1I1
 2381..2386
 /note="putative"
 2407
 /note="putative"
 ORIGIN

Query Match 25.2%; Score 597.4; DB 3; Length 2407;
 Best Local Similarity 81.1%; Pred. No. 2.9e-138;
 Matches 724; Conservative 0; Mismatches 156; Indels 13; Gaps 2;

QY 10 CCATGGATAAGTACGATGTGATTAAAGCCATCGGCAAGGTGCTTCGGGAAAGCATACT 69
 DB 90 CCATGGATAACTTTCACCTGATTAAAGATCATTTGGGAAGGCACCTTTGGGAAGGTGTACT 149
 QY 70 TAGCTAAAGGAATCAGATAGCAAGACATGTGTCTATAAAAGAGATCAATTTTGAAGA 129
 DB 150 TGGCTAAAGATAAATCAAGAACGACGTCACTGTGTCTATAAAAGAAATCAGTTTGACAAAG- 208
 QY 130 TGCCCATACAGAAAGAAAGACGCTTCAAGAAAGAAAGTATTCTTCTGGAAGAAATCAAAAC 189
 DB 209 -----GAAAAGAGCGCTTCAAGAACGAAAGTATTCTTCTGGCTAGATGGAGC 257
 QY 190 ATCCCAACATTTGAGCCTTCTTCAATTCAATTTCAAGAGAAATGCGAGCCTTTTATTGTAA 249
 DB 258 ATCCCAATATCTGAACCTTCTTCACTCGTTTCAAGAGACGGCAGCCTGTTTATTGTAA 317
 QY 250 TGGATATTGATGGAGGGATCTCATGAAGGATCAATAGACACGGGGTGTGTAT 309
 DB 318 TGGATATTGATGGAGGGATCTCATGAGAGGATCCAGAGCAGCGGGGAGTGTAT 377
 QY 310 TTAGTGAAGATCAGATCCTCGGTGTTTGTACAGATTTCTTAGGACTAAACATATTC 369
 DB 378 TCAGCGAAGACAGATCCTGTTGTTGTTTGTACAGATTTCTTAGGACTGAAGCATATTC 437
 QY 370 ATGACAGGAATATTACAGAGGACATATAAGCTCAGACATTTTTTCTTAGCAGAGACG 429
 DB 438 ATGACAGGAATTTTACAGAGGACATATAATCTCAGATATTTTCTTAGCAAGATG 497
 QY 430 GAATGGTGGCAAGCTTGGGACATTTGGTATAGCAAGAGTCTTGAATAATTTCCATGGAAC 489
 DB 498 GAATGGTGGCAAGCTTGGGACATTTGGCAAGCAAGACACTGAATGACTCCATGGAAC 557
 QY 490 TTGCTCGAAGCTTTATTCGAACACCTTACTACTCTGCTCCAGAGATCTGTGAGAAATTAAC 549
 DB 558 TTGCTCAAAACATGTCTCGGACACCTTACTACTCTGCTCCAGAGATCTGTGAGAAACGCG 617
 QY 550 CCTACAAACATAAAGCGATATTTGGTCTTTGGCTGTGCTTATATAGAGCTCTGACAC 609
 DB 618 CATACAAACATAAAGCGACATCTGGTCTTTGGCTGTGCTTATATAGAGCTCTGACAC 677
 QY 610 TTAAACATCTTTTGAGGGTAACAACTTACAGAGCTGGTTTCTGAAAGATTTGTCAAGCAC 669
 DB 678 TCAAGCATCTTTTTCAGAGCAACAACTTCCACCATCTGGTTCTGAAAGATTTGTCAAGGAC 737
 QY 670 ATTTTGCCCAATATCTCGGGGTTTCTCGTGGCTCCATCTTCTGATATCTCAGCTCT 729
 DB 738 GGTGTTGCTCCCAATATCAACCCCACTTCTCTCGTGACCTACAGTCTTGTATACCTCAGCTCT 797
 QY 730 TTAAGATATCTCTCGAGACGACCATCCATAAATTCATTTTGAAGGCGCCCTTTTATG 789
 DB 798 TCAGAGTGTCTCTCAGGACCGGCCATCGGTACGTCCCTTTTGAAGAGACCCCTTTTATG 857
 QY 790 AGAATCTTATTTCCCAATATTTGATCTCTGAGGTC-ATTTCAGGAAGATTCAGTCAATG 848

858 AAACTCTCATTCGCCGATCTCTGTATCTCTGAGGTCTCTTCAAGAAGAAATCAGTCCACG 917
 QY 849 CTTATATCAGACAGAGAGCGCCAGCTTCTCGACATCTCGGGAAGGTGGTCCA 901
 DB 918 CTCACATGGAGAATCGCCATTCGCCCCACAGCTTGTGGAGAGTTAGTCCA 970

RESULT 6
 AY405588
 LOCUS
 DEFINITION
 Homo sapiens HCM2274 gene, VIRTUAL TRANSCRIPT, partial sequence,
 genomic survey sequence.
 ACCESSION
 AY405588
 VERSION
 AY405588.1
 GI:39761562
 KEYWORDS
 GSS.
 SOURCE
 Homo sapiens (human)
 ORGANISM
 Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 REFERENCE
 1 (bases 1 to 603)
 Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejarival,A.,
 Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,
 Ferriera,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J.,
 Adams,M.D. and Gargill,M.
 Inferring nonneutral evolution from human-chimp-mouse orthologous
 gene trios
 Science 302 (5652), 1960-1963 (2003)
 JOURNAL
 PUBMED
 14671302
 REFERENCE
 2 (bases 1 to 603)
 Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejarival,A.,
 Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,
 Ferriera,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J.,
 Adams,M.D. and Gargill,M.
 Direct Submisison
 Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive,
 Rockville, MD 20850, USA
 COMMENT
 This sequence was made by sequencing genomic exons and ordering
 them based on alignment.
 FEATURES
 Location/Qualifiers
 1..603
 /organism="Homo sapiens"
 /mol_type="genomic DNA"
 /db_xref="taxon:9606"
 <1..>603
 /locus_tag="HCM2274"
 ORIGIN

Query Match 25.1%; Score 594; DB 9; Length 603;
 Best Local Similarity 99.2%; Pred. No. 1.5e-137;
 Matches 597; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 231 GCGAGCTGTTTATTGTAATGGAATATTGTATGAGGGGATCTCATGAAAGGATCAAT 290
 DB 1 GCGAGCTGTTTATTGTAATGGAATATTGTATGAGGGGATCTCATGAAAGGATCAAT 60
 QY 291 AGACAAACGGGGTGTGTTATTTAGTGAAGATCAGATCCCTCGGTGGTTGTACAGATTTCT 350
 DB 61 AGACAAACGGGGTGTGTTATTTAGTGAAGATCAGATCCCTCGGTGGTTGTACAGATTTCT 120
 QY 351 CTAGGACTTAAACATATTCATGACAGGAGATATTACACAGGACATATAAGCTCAGAAC 410
 DB 121 CTAGGACTTAAACATATTCATGACAGGAGATATTACACAGGACATATAAGCTCAGAAC 180
 QY 411 ATTTTCTTTAGCAAGAAACGGAATGTTGGCAAGCTTTGGGACTTTGGTATAGCAAGATC 470
 DB 181 ATTTTCTTTAGCAAGAAACGGAATGTTGGCAAGCTTTGGGACTTTGGTATAGCAAGATC 240
 QY 471 CTGAATAATTCATGGAACCTTCTGAACTTGTATTGGAAACACCTTACTCTGTCCTCCCA 530
 DB 241 CTGAATAATTCATGGAACCTTCTGAACTTGTATTGGAAACACCTTACTCTGTCCTCCCA 300
 QY 531 GAGATCTGTGAGATAAACCCTTACAAATAAAGCGATATTTGGTCTCTTTGGTGTGTCT 590

```

Db      301  |||||GAGATCTGTGAGAAATTAACCCCTACACAAATAAAACGGAATATTGTCCTCTCGCTGTGTC 360
Qy      591  |||||TTATATGAGCTCTGCACACTTAAACATCTCTTTTGGGGTAAACAACTTACAGCAGCTGGTT 650
Db      361  |||||TTATATGAGCTCTGCACACTTAAACATCTCTTTTGGGGTAAACAACTTACAGCAGCTGGTT 420
Qy      651  |||||CTGAAGATTGTCAAGCACAATTTTCCCAATATCTCTCGGGGTTTCTCGTGAGCTCCAT 710
Db      421  |||||CTGAAGATTGTCAAGCACAATTTTCCCAATATCTCTCGGGGTTTCTCGTGAGCTCCAT 480
Qy      711  |||||TCCTTGATATCTCAGCTCTTCAAGTATCTCTCGAGACCGACCATCCATAAATTCATTT 770
Db      481  |||||TCCTTGATATCTCAGCTCTTCAAGTATCTCTCGAGACCGACCATCCATAAATTCATTT 540
Qy      771  |||||TTGAAAGGCCCTTTTATAGAAATCTTATTCCCAATATTTGACTCTCTGAGGTCATTCAG 830
Db      541  |||||TTGAAAGGCCCTTTTATAGAAATCTTATTCCCAATATTTGACTCTCTGAGGTAAGTTT 600
Qy      831  GA 832
Db      601  GA 602

```

RESULT 7

```

BX417448      872 bp  mRNA  linear  EST 01-MAY-2004
LOCUS      BX417448 Homo sapiens PLACENTA Homo sapiens cDNA clone CS0DE009YJ07
DEFINITION  5-PRIME, mRNA sequence.
ACCESSION  BX417448
VERSION    BX417448.2 GI:46926799
KEYWORDS   EST.
SOURCE     Homo sapiens (human)
ORGANISM   Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
            1 (bases 1 to 872)
            Li, W.B., Gruber, C., Jessee, J. and Polayes, D.
            Full-length cDNA libraries and normalization
            Unpublished (2001)
            On May 13, 2003 this sequence version replaced gi:30642067.
            Contact: Genoscope
            Genoscope - Centre National de Sequencage
            BP 191 91006 Evry cedex - France
            Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
            1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime
            end enriched, double-strand cDNA was digested with Not I and cloned
            into the Not I and EcoRV sites of the pCMVSPORT 6 vector. Library
            was not normalized. Library was constructed by Life Technologies, a
            division of Invitrogen.
            This sequence belongs to sequence cluster 6942.r
            There is a virtual cDNA representing this cluster. For more
            information about this cluster and the virtual cDNA, see
            http://www.genoscope.cns.fr/cdnasr=CS0DE009CE04QP&c=6942.r.
            Location/Qualifiers
            1..872
            /organism="Homo sapiens"
            /mol_type="mRNA"
            /db_xref="taxon:9606"
            /clone="CS0DE009YJ07"
            /tissue_type="PLACENTA"
            /note="Vector: pCMVSPORT 6; 1st strand cDNA was primed
            with a NotI-oligo(dT) primer. Five prime end enriched,
            double-strand cDNA was digested with Not I and cloned into
            the Not I and EcoRV sites of the pCMVSPORT 6 vector.
            Library was not normalized."

```

FEATURES

```

source
1..872
CD700193      670 bp  mRNA  linear  EST 25-JUN-2003
DEFINITION  EST16717 human nasopharynx Homo sapiens cDNA, mRNA sequence.
ACCESSION  CD700193
VERSION    CD700193.1 GI:32230189
KEYWORDS   EST.
SOURCE     Homo sapiens (human)
ORGANISM   Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
            1 (bases 1 to 670)
            Liu, X.-Q., Zhou, Y., Zhang, L.-J., Xu, H., Chen, H.-K., Pan, Z.-G. and
            Zeng, Y.-X.
            Transcriptional Gene Expression Profile of Human Nasopharynx
            Unpublished (2003)
            Query Match      24.3%; Score 574.8; DB 5; Length 872;
            Best Local Similarity 84.1%; Pred. No. 1.1e-132;
            Matches 695; Conservative 1; Mismatches 123; Indels 7; Gaps 4;

```

ORIGIN

```

1..872
CD700193      670 bp  mRNA  linear  EST 25-JUN-2003
DEFINITION  EST16717 human nasopharynx Homo sapiens cDNA, mRNA sequence.
ACCESSION  CD700193
VERSION    CD700193.1 GI:32230189
KEYWORDS   EST.
SOURCE     Homo sapiens (human)
ORGANISM   Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
            1 (bases 1 to 670)
            Liu, X.-Q., Zhou, Y., Zhang, L.-J., Xu, H., Chen, H.-K., Pan, Z.-G. and
            Zeng, Y.-X.
            Transcriptional Gene Expression Profile of Human Nasopharynx
            Unpublished (2003)

```

COMMENT

Contact: Yixin Zeng
Cancer Center
Sun Yat-sen University
651 Dongfeng Road East, Guangzhou 510060, China
Tel: 86-1380-9770-743
Fax: 86-20-8775-4506
Email: yxzeng@gzsums.edu.cn

FEATURES

source
1..670
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/tissue_type="normal nasopharynx"
/clone_lib="human nasopharynx"
/note="ESTs generated from a normal nasopharynx cDNA library from southern Chinese"

ORIGIN

Query Match 23.1%; Score 547.8; DB 6; Length 670;
Best Local Similarity 99.6%; Pred. No. 5.9e-126;
Matches 549; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 CATTGGAGACCATGATAAGTACCATGTGATTAGGCCATCGGCCAAGGTGCCTTCGGGA 60
DB 105 CATTGGAGACCATGATAAGTACCATGTGATTAGGCCATCGGCCAAGGTGCCTTCGGGA 164
QY 61 AAGCATCTTAGCTAAAGGGAAATCAGATAGCAAGCACTGTGTGCATAAAGAGATCAATT 120
DB 165 AAGCATCTTAGCTAAAGGGAAATCAGATAGCAAGCACTGTGTGCATAAAGAGATCAATT 224
QY 121 TTGAAGAATGCCATACAGAAAGAAAGAGCTTCAAGAGAAAGAGTCAATCTTCGAAA 180
DB 225 TTGAAGAATGCCATACAGAAAGAAAGAGCTTCAAGAGAAAGAGTCAATCTTCGAAA 284
QY 181 AGATGAACATCCCAACATTTAGCTCTTCTCAATTCATTTCAAGAAATGGCAGGCTGT 240
DB 285 AGATGAACATCCCAACATTTAGCTCTTCTCAATTCATTTCAAGAAATGGCAGGCTGT 344
QY 241 TTATGTAAATCGAATATGTGATCGAGGGATCTCATGAAGAGTCAATAGCAACGGG 300
DB 345 TTATGTAAATCGAATATTTGTATGGGGGATCTCATGAAGAGTCAATAGCAACGGG 404
QY 301 GTGTGTTATTTAGTGAAGATCAGATCCTCGGTGGTTGTACAGATTTCTTAGGACTAA 360
DB 405 GTGTGTTATTTAGTGAAGATCAGATCCTCGGTGGTTGTACAGATTTCTTAGGACTAA 464
QY 361 AACATATTCATGACAGGAAGATATTACAGGGACATATAAGCTCAGAACATTTTCTTA 420
DB 465 AACATATTCATGACAGGAAGATATTACAGGGACATATAAGCTCAGAACATTTTCTTA 524
QY 421 GCAAGAACGGAATGTGGCAAGCTTGGGACTTTGGTATAGCAAGAGTCTCTGAATAAT 480
DB 525 GCAAGAACGGAATGTGGCAAGCTTGGGACTTTGGTATAGCAAGAGTCTCTGAATAAT 584
QY 481 CCATGGAACTTGCTCGAACTTGTATATGGAACACCTTACTACCTGTCCCCAGAGATCTGTC 540
DB 585 CCATGGAACTTGCTCGAACTTGTATATGGAACACCTTACTACCTGTCCCCAGAGATCTGTC 644
QY 541 AGAATAACCC 551
DB 645 AGAATAACCC 655

RESULT 9

BQ917635
LOCUS BQ917635 928 bp mRNA linear EST 20-AUG-2002
DEFINITION AGENCOURT_8888753 NIH_MGC_129 Mus musculus cDNA clone IMAGE:6311513
5', mRNA sequence.
ACCESSION BQ917635
VERSION BQ917635.1 GI:22332333
KEYWORDS EST.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus

REFERENCE

1 (bases 1 to 928)

AUTHORS

NIH-MGC <http://mgc.nci.nih.gov/>.

TITLE

National Institutes of Health, Mammalian Gene Collection (MGC)

JOURNAL

Unpublished (1999)

COMMENT

Contact: Robert Strausberg, Ph.D.

Email: cgapbs-remail.nih.gov

Tissue Procurement: Susan L. Sullivan, PhD.

cDNA Library Preparation: ResGen, Invitrogen Corp

cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

DNA Sequencing by: Agencourt Bioscience Corporation

Clone Distribution: MGC clone distribution information can be

found through the I.M.A.G.E. Consortium/LLNL at:

<http://image.llnl.gov>

Plate: LLAM13735 row: i column: 18

High quality sequence stop: 734.

FEATURES

Location/Qualifiers

1..928

/organism="Mus musculus"

/mol_type="mRNA"

/db_xref="taxon:10090"

/clone="IMAGE:6311513"

/lab_host="DH10B (phage-resistant)"

/clone_lib="NIH MGC 129"

/note="Organ: olfactory epithelium; Vector: pCMV-SPORT6.1;

Site_1: EcoRV; Site_2: NotI; Cloned unidirectionally.

Primer: Oligo dt. Average insert size 2.2 kb. Constructed

by ResGen, Invitrogen Corp. Note: this is a NIH_MGC

Library."

ORIGIN

Query Match 21.8%; Score 516.8; DB 5; Length 928;
Best Local Similarity 81.2%; Pred. No. 3.8e-118; Indels 16; Gaps 3;
Matches 642; Conservative 0; Mismatches 133; Indels 16; Gaps 3;

QY 10 CCATGGATAAGTACCATGTGATTAGGCCATCGGCCAAGGTGCCTTCGGGAAAGCACTACT 69
DB 91 CCATGGATAAGTACCATGTGATTAGGCCATCGGCCAAGGTGCCTTCGGGAAAGGTACT 150
QY 70 TAGCTAAAGGGAATCAGATAGCAAGCACTGTGTGCATAAAGAGATCAATTTGAAAGA 129
DB 151 TGCTTAAAGATAAATCAGAAAGCAGTCACTGTGTGCATAAAGAGATCAATTTGCAAG- 209
QY 130 TGCCCATACAGAAAGAAAGAGCTTCAAGAGAAAGAGTGAATCTTCTGGAAGAGATGAAC 189
DB 210 -----GAAAAAGAGGCTTCAAGAGAGAGTGAATCTTCTGCTAGGATGGAGC 258
QY 190 ATCCCAACATTTAGCTTCTCAATTCATTTCAAGAGAGTGGCAGGCTGTTATTGTAA 249
DB 259 ATCCCAATATCGTAACCTTCTTCAGCTCGTTTCAAGAGAGCGGAGGCTGTTATTGTAA 318
QY 250 TGGNATATTGTATGATGAGGGGATCTCATGAAAGAGATCAATAGACAACGGGTGTGTTAT 309
DB 319 TGGNATATTGTATGATGAGGGGATCTCATGAGAGAGTCCAGAGGAGCGGGAGTGTATG 378
QY 310 TTAGTGAAGATCAGATCCTCGGTGGTTGTGTACAGATTTCTTAGGACTAAACATATTC 369
DB 379 TCAGCGAAGAGCCAGATCCTGTGTTGTTGTACAGATTTCTTAGGACTGAAGCATATTC 438
QY 370 ATCAGAGGAGATATTACAGAGGAGCATAAAGCTCAGAACATTTTCTTAGCAGAACG 429
DB 439 ATGACAGGAAGATTTTACAGAGGAGCATAAATCTCAGANATATTTTCTTAGCAGAGATG 498
QY 430 GAATGGTGGCAAGCTTGGGACTTTGGTATAGCAAGAGTCTCGAATAATTTCCATGGAAC 489
DB 499 GAATGGTGGCAAGCTTGGGACTTTGGAAACAGCAAGAAACACTGAATGACTCCATGGAAC 558
QY 490 TTGCTCGAAGCTTGTATTGGAACACCTTACTACCTGTGCCAGAGATCTGTGAGATTAAC 549
DB 559 TTGCTCAAAACATGTCTCTGGGACACCTTACTACCTGTGCCAGAGATCTGTGAGAACGC 618
QY 550 CCTACAACAATAAAACGGGATATTTGGTCTCTTGGCTGTGTCTTATATAGAGCTCTGCACAC 609

```

Db      619 CATACACATAAAACCGGACATCTGGTCTCTGGCTGTGTCTATATGAGCTCTGCACAC 678
QY      610 TTAACATCTTTTGGGGTAAACAATTTACAGCAGCTGGTCTGGAAGATTGTCAAGCAC 669
Db      679 TCAAGCATCTTTTGGAGAGCAACAATCTCCACCATCTGGTCTGAAGATTGTCAAGGAC 738
QY      670 ATTTTCCCAATATCTCCGGGTTTCTCGTAGCTCCATTC--TTGATATCTCAGCT 727
Db      739 GTTGTCTCCATATCACCCCACTTCTCTCGTAGCTACAGTNCCTTGATACCTCAGCT 798
QY      728 CTTTCAAGTATCTCTCCGAGACC--GACCATCCATAAATTCATTTTGAAGAGCCCTTT 785
Db      799 CTTCAAGTGTCTCTCAGAGCCGCGCATCGTTACGTCCCTTTTGAAGAGCCCTTT 858
QY      786 TTAGAGAACT 796
Db      859 TTTAGAAACT 869

CD243608      856 bp mRNA linear EST 22-MAY-2003
AGENCOURT_14121106 NIH_MGC_180 Homo sapiens cDNA clone
IMAGE:30383658 5', mRNA sequence.
CD243608
EST.
CD243608.1 GI:31004072
Homo sapiens (human)
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
NIH-MGC http://mgs.nci.nih.gov/
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgabs-r@mail.nih.gov
Tissue Procurement: Dr. Michael Brownstein
cDNA Library Preparation: Invitrogen Corp
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: NDAM447 row: n column: 19
High quality sequence stop: 545.
Location/Qualifiers
1..856
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:30383658"
/lab_host="PH108-Tcn A ( T1 and T5 phage resistances)"
/clone_lib="NIH_MGC_180"
Note="Organ: Testis; Vector: pCMV-SPORT6.1; Site.1: NotI;
Site.2: EcoRV (destroyed); Library is oligo-dT primed and
directionally cloned (EcoRV site is destroyed upon
cloning). Average insert size 1.68 kb. Library was
constructed by (Invitrogen). Note: this is a NIH_MGC
Library."

ORIGIN
Query Match 20.5%; Score 485; DB 6; Length 856;
Best Local Similarity 94.3%; Pred. No. 3.5e-110;
Matches 548; Conservative 0; Mismatches 25; Indels 8; Gaps 4;

QY      1 CATTGAGACCATGATAGTACGATGATTAAAGCCATCGGCAAGTGCCTTCGGGA 60
Db      137 CATTGAGACCATGATAGTACGATGATTAAAGCCATCGGCAAGTGCCTTCGGGA 196
QY      61 AAGCATCTTAGCTAAAGGAAATCAGATAGCAAGCACTGTGTCTATAAAGAGATCAATT 120

```

```

Db      197 AAGCATCTTAGCTAAAGGAAATCAGATAGCAAGCACTGTGTCTATAAAGAGATCAATT 256
QY      121 TTGAAAAGATGCCCATACAGAAAAAGAGCTTCAAAAGAAAGAGTCAATCTTCTCGAAA 180
Db      257 TTGAAAAGATGCCCATACAGAAAAAGAGCTTCAAAAGAAAGAGTCAATCTTCTCGAAA 316
QY      181 AGATGAACATCCCAACATTTGTAGCCTTCTCAATTCAATTTCAAGAGAATGGCAGGCTGT 240
Db      317 AGATGAACATCCCAACATTTGTAGCCTTCTCAATTCAATTTCAAGAGAATGGCAGGCTGT 376
QY      241 TTATTGTAATGGAAATATTGTGATGGAGGGATCTCATGAAAGAGATCAATAGACAAACGGG 300
Db      377 TTATTGTAATGGAAATATTGTGATGGAGGGATCTCATGAAAGAGATCAATAGACAAACGGG 436
QY      301 GTGTGTTATTAGTGAAGATCAGATCCTCGTGGTGGTTGTACAGATTCTCTAGGACTAA 360
Db      437 GTGTGTTATTAGTGAAGATCAGATCCTCGTGGTGGTTGTACAGATTCTCTAGGACTAA 496
QY      361 AACATATTATGACAGGAAGATATTACACAGGGACATAAAGCTCAGAACATTTTCTTCTTA 420
Db      497 AACATATTATGACAGGAAGATATTACACAGGGACATAAAGCTCAGAACATTTTCTTCTTA 556
QY      421 GCAAGAAACGGAATGGTGGCAAAAGCTTGGGACATTGGTATAGCAAGAGT--CCTGAATAAT 479
Db      557 GTCAGAACGGAATGGTGGCAAAAGCTTGGGACATTGGTATAGCAAGAGTCCCTGAATAAT 616
QY      480 TCCAT--GGAATCTCTCGAACTTGTATTGG--AACACCTTACTACCTGTGCCCC--AGA 532
Db      617 TCCATGGGAACTTGTCTGAACTTGTATTGGGAAACACCTTTACTACCTGTGCCCCCAGAAGA 676
QY      533 GATCTGTGAGATAAACCCTACAACTAAACGATATTT 573
Db      677 TCTGTTCAGATAAACCCTACAACTAAACGAGGATAT 717

RESULT 11
LOCUS      AY405589
DEFINITION Pan troglodytes HCM2274 gene, VIRTUAL TRANSCRIPT, partial sequence,
genomic survey sequence.
ACCESSION AY405589
VERSION AY405589.1 GI:39761563
KEYWORDS GSS.
SOURCE      Pan troglodytes (chimpanzee)
ORGANISM    Pan troglodytes
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Pan.
REFERENCE   1 (bases 1 to 603)
AUTHORS     Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejarawal,A.,
            Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,
            Ferreira,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J.,
            Adams,M.D. and Cargill,M.
            Inferring nonneutral evolution from human-chimp-mouse orthologous
            gene trios
JOURNAL     Science 302 (5652), 1960-1963 (2003)
PUBMED     14671302
REFERENCE   2 (bases 1 to 603)
AUTHORS     Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejarawal,A.,
            Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,
            Ferreira,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J.,
            Adams,M.D. and Cargill,M.
            Direct Submission
            Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive,
            Rockville, MD 20850, USA
            This sequence was made by sequencing genomic exons and ordering
            them based on alignment.
FEATURES     Location/Qualifiers
             source
               1..603
               /organism="Pan troglodytes"
               /mol_type="genomic DNA"
               /db_xref="taxon:9598"
               <1..>603
               /locus_tag="HCM2274"
             gene

```



```

ORIGIN
Query Match      19.6%; Score 464.8; DB 9; Length 603;
Best Local Similarity 77.9%; Pred. No. 3.7e-105;
Matches 469; Conservative 0; Mismatches 133; Indels 0; Gaps 0;

QY 231 GCGAGGCTGTTTATTGTAATGGAATATTGATGAGGGGATCTCATGAAAAGGATCAAT 290
DB 1 GCGAGGCTGTTTATTGTAATGGAATATTGATGAGGGGATCTCATGAAAAGGATCAAT 60

QY 291 AGACAACGGGGTGTGTTATTATGTAAGATCAGATCCTCGGGTGGTTGTTGACAGATTCT 350
DB 61 AGACAACGGGGTGTGTTATTATGTAAGATCANNNNCTCGGGTGGTTGTTGACAGATTCT 120

QY 351 CTAGAGCTAAACATATTCTATGACAGAGAGATATTACACAGGGGACATAAAGCTCAGAAC 410
DB 121 CTAGAGCTAAACATATTCTATGACAGAGAGATATTACACAGGGGACATAAAGCTCAGAAC 180

QY 411 ATTTTCTTAGCAAGAACGGAATGTTGGAAGCTTGGGGACTTTGGGTATAGCAAGAGTC 470
DB 181 ATTTTCTTAGCAAGAACGGAATGTTGGAAGCTTGGGGACTTTGGGAACAGCAAGAACA 240

QY 471 CTGAATATTCTAGGAACTTCTCGAACTTGTATGGAACCTTACTACTGTCCTCCCA 530
DB 241 CTGAATGACTCCATGAACTTCTCGAACTTGTATGGAACCTTACTACTGTCCTCCCA 300

QY 531 GAGATCTGTGAGAAATAAACCCCTACAACTAAACGGAATTTGGTCTCTTTGGGTGTGTC 590
DB 301 GAGATCTGTGAGAACTTCTCGAACTTGTATGGAACCTTACTACTGTCCTTTGGGTGTGTC 360

QY 591 TTATATGAGCTCTGACACCTTAAACATCTTTTGAAGGTAACAACCTTACAGCAGCTGGTT 650
DB 361 TTATATGAGCTCTGACACCTTAAACATCTTTTGAAGGTAACAACCTTACAGCAGCTGGTT 420

QY 651 CTGAAGATTGTCAGAGCATTGTCGCCCAATATCTCGGGGTTTCTCGTGAGCTCCAT 710
DB 421 CTGAAGATTGTCAGAGCATTGTCGCCCAATATCTCGGGGTTTCTCGTGAGCTCCAT 480

QY 711 TCCTTGATATCTCAGCTCTTTCAAGTATCTCTCGAGACCGGACCATCCATAAATCCATT 770
DB 481 TCCTTGATATCTCAGCTCTTTCAAGTATCTCTCGAGACCGGACCATCCATAAATCCATT 540

QY 771 TTGAAAGGCCCTTTTAGAGAACTTTATCCCAATATTGATCTCGAGGTCATTTCAG 830
DB 541 TTGAAAGGCCCTTTTAGAGAACTTTATCCCAATATTGATCTCGAGGTAAGTTT 600

QY 831 GA 832
DB 601 GA 602

RESULT 12
AY405590 603 bp DNA linear GSS 12-DEC-2003
LOCUS Mus musculus HCM2274 gene, VIRTUAL TRANSCRIPT, partial sequence,
DEFINITION genomic survey sequence.
ACCESSION AY405590
VERSION AY405590.1 GI:39761564
KEYWORDS GSS.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 603)
Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejariwal,A.,
Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,
Ferriera,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J.,
Adams,M.D. and Cargill,M.
Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive,
Rockville, MD 20850, USA
COMMENT This sequence was made by sequencing genomic exons and ordering
them based on alignment.
FEATURES
source Location/Qualifiers
1..603
/organism="Mus musculus"
/mol_type="genomic DNA"
/db_xref="taxon:10090"
<1..>603
/locus_tag="HCM2274"

ORIGIN
Query Match      19.2%; Score 455.4; DB 9; Length 603;
Best Local Similarity 84.9%; Pred. No. 8.5e-103;
Matches 510; Conservative 0; Mismatches 91; Indels 0; Gaps 0;

QY 231 GCGAGGCTGTTTATTGTAATGGAATATTGATGAGGGGATCTCATGAAAAGGATCAAT 290
DB 1 GCGAGGCTGTTTATTGTAATGGAATATTGATGAGGGGATCTCATGAGAGGATCCAG 60

QY 291 AGACAACGGGGTGTGTTATTATGTAAGATCAGATCCTCGGGTGGTTGTTGACAGATTCT 350
DB 61 AGGACGCGGGGATGATGTTTCAGGAGAACGAGATCCTGTGTGTTGTTGACAGATTCT 120

QY 351 CTAGAGCTAAACATATTCTATGACAGAGAGATATTACACAGGGGACATAAAGCTCAGAAC 410
DB 121 CTAGAGCTAAACATATTCTATGACAGAGAGATATTACACAGGGGACATAAATCTCAGAAT 180

QY 411 ATTTTCTTAGCAAGAACGGAATGTTGGAAGCTTGGGGACTTTGGGTATAGCAAGAGTC 470
DB 181 ATTTTCTTAGCAAGAACGGAATGTTGGAAGCTTGGGGACTTTGGGAACAGCAAGAACA 240

QY 471 CTGAATATTCTAGGAACTTCTCGAACTTGTATGGAACCTTACTACTGTCCTCCCA 530
DB 241 CTGAATGACTCCATGAACTTCTCGAACTTGTATGGAACCTTACTACTGTCCTCCCA 300

QY 531 GAGATCTGTGAGAAATAAACCCCTACAACTAAACGGAATTTGGTCTCTTTGGGTGTGTC 590
DB 301 GAGATCTGTGAGAACTTCTCGAACTTGTATGGAACCTTACTACTGTCCTTTGGGTGTGTC 360

QY 591 TTATATGAGCTCTGACACCTTAAACATCTTTTGAAGGTAACAACCTTACAGCAGCTGGTT 650
DB 361 TTATATGAGCTCTGACACCTTAAACATCTTTTGAAGGTAACAACCTTACAGCAGCTGGTT 420

QY 651 CTGAAGATTGTCAGAGCATTGTCGCCCAATATCTCGGGGTTTCTCGTGAGCTCCAT 710
DB 421 CTGAAGATTGTCAGAGCATTGTCGCCCAATATCTCGGGGTTTCTCGTGAGCTCCAT 480

QY 711 TCCTTGATATCTCAGCTCTTTCAAGTATCTCTCGAGACCGGACCATCCATAAATCCATT 770
DB 481 TCCTTGATATCTCAGCTCTTTCAAGTATCTCTCGAGACCGGACCATCCATAAATCCATT 540

QY 771 TTGAAAGGCCCTTTTAGAGAACTTTATCCCAATATTGATCTCGAGGTCATTTCAG 830
DB 541 TTGAAAGGCCCTTTTAGAGAACTTTATCCCAATATTGATCTCGAGGTAAGTTT 600

QY 831 G 831
DB 601 G 601

RESULT 13
BB666328 659 bp mRNA linear EST 26-OCT-2001
LOCUS BB666328 RIKEN full-length enriched, 2 days pregnant adult female
DEFINITION ovary Mus musculus cDNA clone E330012F19 5', mRNA sequence.
ACCESSION BB666328

```


into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library was normalized. Library was constructed by Life Technologies, a division of Invitrogen. This sequence belongs to sequence cluster 6942.r, and it belongs to a clone representative of this cluster. For more information about this cluster and the virtual cDNA, see <http://www.genoscope.cns.fr/cdna?ss=CSOBAG0542E01.CS05093> 15c=6942.r

FEATURES

Location/Qualifiers

```

1. 666678.1 Homo sapiens
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0DL005YK18"
/cell_type="B CELLS (RAMOS CELL LINE) COT 25-NORMALIZED"
/cell_line="RAMOS CELL LINE"
/clone_lib="Homo sapiens B CELLS (RAMOS CELL LINE) COT
25-NORMALIZED"
/notes="1st strand cDNA was primed with a NotI-oligo(dT)
primer. Five prime end enriched, double-strand cDNA was
digested with Not I and cloned into the Not I and EcoR V
sites of the pCMVSPORT 6 vector. Library was normalized."

```

ORIGIN

Query Match	18.8%;	Score	446.2;	DB	5;	Length	678;
Best Local Similarity	99.3%;	Prod. No.	1.8e-100;				
Matches	448;	Conservative	0;	Mismatches	3;	Indels	0;
Gaps	0;						
770	TTTGAAGAGGCCCTTTT	TAGAGAACTCTTAT	TCCCAATATTTG	ACTCTCTGAGGTC	CAATCA	829	
45	TTTGAAGAGGCCCTTTT	TAGAGAACTCTTAT	TCCCAATATTTG	ACTCTCTGAGGTC	CAATCA	104	
830	GGAAGAAATTCAGTCA	CATGCTTTATATG	CAGAGCAGGAGCG	CGAGCTTCTCGACATG	CTGG	889	
105	GGAAGAAATTCAGTCA	CATGCTTTATATG	CAGAGCAGGAGCG	CGAGCTTCTCGACATG	CTGG	164	
890	GAAAGTGTCTCAGAAG	TGTAAAAATCAAAA	AGTGTAGATTCCGGG	GAAGAGTCCCA	CCCAAG	949	
165	GAAAGTGTCTCAGAAG	TGTAAAAATCAAAA	AGTGTAGATTCCGGG	GAAGAGTCCCA	CCCAAG	224	
950	ATCAAGGATATCTGTG	CCAAATPAAAGAA	TGCTATATTG	CATAGAAATGAAT	GAGAGACC	1009	
225	ATCAAGGATATCTGTG	CCAAATPAAAGAA	TGCTATATTG	CATAGAAATGAAT	GAGAGACC	284	
1010	ACCAGCTGGAGCCCA	GAGGCCAGATCTAT	AAAAATGATAGAA	AGACCCAAAAT	TGCTGC	1069	
285	ACCAGCTGGAGCCCA	GAGGCCAGATCTAT	AAAAATGATAGAA	AGACCCAAAAT	TGCTGC	344	
1070	TGCTGTGGACANTTA	GAATTTATTTATG	CTCAACTTGAT	TGCTAGGAGGAGAG	CCCA	1129	
345	TGCTGTGGACANTTA	GAATTTATTTATG	CTCAACTTGAT	TGCTAGGAGGAGAG	CCCA	404	
1130	CAAPACCAAGTTATCA	CCCTATTCCTCA	GAAAAATACTG	SAGTTAGAGGATTAC	GGTCA	1189	
405	CAAPACCAAGTTATCA	CCCTATTCCTCA	GAAAAATACTG	SAGTTAGAGGATTAC	GGTCA	464	
1190	AACGAGGCATGGTCC	ATCCCCCAAGTCA	ATGG	1220			
455	AACGAGGCATGGTCC	ATCCCCCAAGTCA	ATGG	495			

RESULT 15

B8623078		B8623078	639 bp	mRNA	linear	EST 26-OCT-2001
LOCUS		B8623078 RIKEN full-length enriched, 10 days neonate cerebellum Mus musculus cDNA clone G5304UJ22 5'				
DEFINITION		Mus musculus				
ACCESSION		B8623078.l	G1:16461881			
VERSION		EST.				
KEYWORDS		Mus musculus (house mouse)				
SOURCE		Mus musculus				
ORGANISM		Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;				
		Mammalia; Euarchia; Rodentia; Sciurognathi; Muridae; Murinae; Mus;				

BamHI "

ORIGIN

```
Query Match      16.5%; Score 392; DB 2; Length 639;
Best Local Similarity 82.7%; Pred. No. 6.8e-87;
Matches 465; Conservative 0; Mismatches 55; Indels 12; Gaps 1;

Qy 10 CCATGGATAAGTACGATGTGATTAAGGCCATCGGCAAGGTGCGCTTCGGGAAAGCATACT 69
Db 90 CCATGGATAACTTCACTGATTAGATCATTTGGGGAAGGCACCTTTGGGAAGGTGCTACT 149
Qy 70 TAGCTAAGGGAATCAGATAGCAGACACTGTCTCATAAAGAGATCAATTTTGAAGAAGA 129
Db 150 TGGCTAAAGTAATCAGAAGCAGTCACCTGTCTATAAAGAAATCAGTTTGACAAAG- 208
Qy 130 TGCCCATACAGAAAGAAAGAGCTTCAAGAAAGAAAGATGATTCTTCTGGAAGAAATGAAAC 189
Db 209 -----GAAAGAGGCGCTCAAGAAACGAGTGTCTTCTGGCTAGGATGGAGC 257
Qy 190 ATCCCAACATGTAGCCTTCTTCAATTTCATTTCAAGAGATGCGCAGGCTGTTTATTGTAA 249
Db 258 ATCCCAATATCGTAACCTTCTTCAAGTCTGTTTCAAGAGAACGCGCAGGCTGTTTATTGTAA 317
Qy 250 TGGAAATATCTGATGGAGGGGATCTCATGAAAAGGATCAATAGACAAACGGGGTGTGTAT 309
Db 318 TGGATACTGTGATGGAGGGGATCTCATGAGAGGATCCAGAGGCGAGGGAGTGATGT 377
Qy 310 TTAGTGAAGATCAGATCCTCGGTGGTTTGTACAGATTTCTTAGGACTAAACATATTTC 369
Db 378 TCAGCGAAGACCAAGATCCTGTGTGGTTTGTACAGATTTCTTAGGACTGAAGCATATTTC 437
Qy 370 ATGACAGGAGATATTACAGAGGACATAAAGCTCAGAACATTTTCTTAGCAAGAACG 429
Db 438 ATGACAGGAGATTTTACAGAGGACATAAAATCTCAGAAATATTTTCTTAGCAAGATG 497
Qy 430 GAATGTGTGCAAGCTTGGGGACTTTGGTATAGCAAGAGTCTCTGAATAATTCCATGGAAC 489
Db 498 GAATGTGTGCAAGCTCGGGGACTTTGGAACACCNAGAACACTGATGACTCCATGGAAC 557
Qy 490 TTGCTCGAATTTGTATTGGAAACCTTACTACTCTGCTCCCGAGAGATCTGTCTAGAAATAAC 549
Db 558 TTGCTCAACATGTGTGGGACACCTTACTACTCTGCTCCCGAAGATCTGGCAGAACAGGC 617
Qy 550 CCTACAAACATAAAACGGATAT 571
Db 618 CATACAAACATAAAACGGACAT 639
```

Search completed: November 8, 2004, 01:59:21
Job time : 6685 secs

This Page Blank (uspto)